SMTCoq: skeptical cooperation between SAT/SMT solvers and Coq

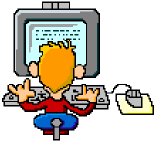
Michaël Armand Germain Faure Benjamin Grégoire
Chantal Keller Laurent Théry Benjamin Werner

INRIA – École Polytechnique

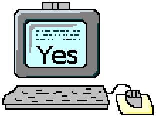
March, 30th 2013

Motivation (1/2)

COQ

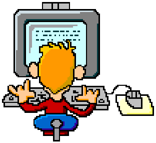


AUTOMATIC THEOREM PROVER

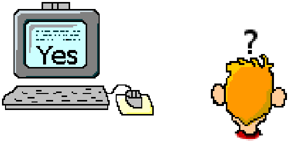


Motivation (1/2)

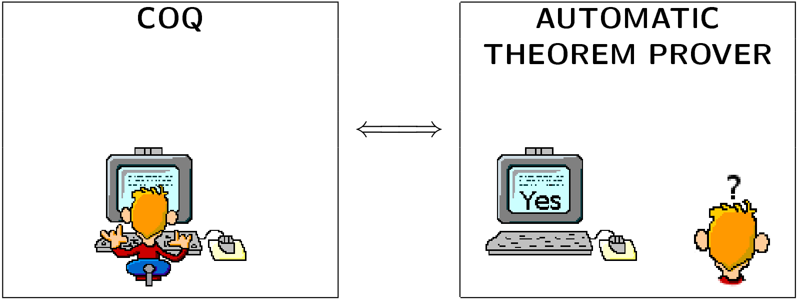
COQ



AUTOMATIC THEOREM PROVER



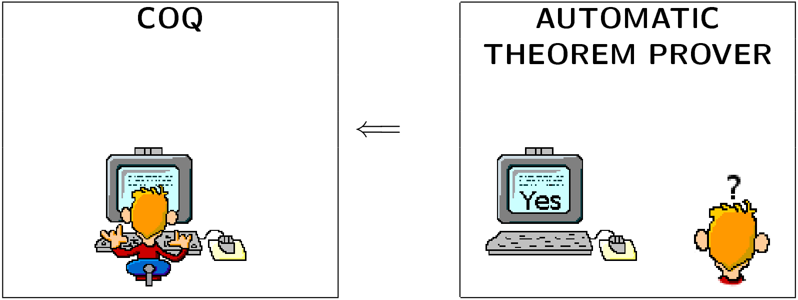
Motivation (1/2)



COQ

AUTOMATIC THEOREM PROVER

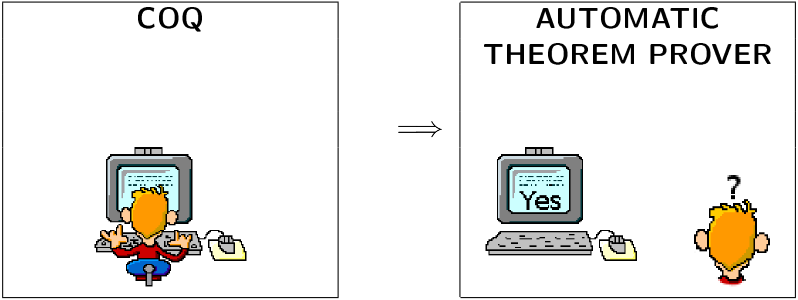
Motivation (1/2)



COQ ⟸ AUTOMATIC THEOREM PROVER

AUTOMATION

Motivation (1/2)



COQ

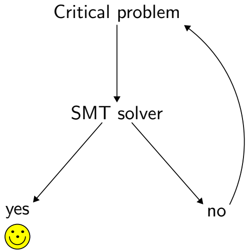
$\Longrightarrow$

AUTOMATIC THEOREM PROVER

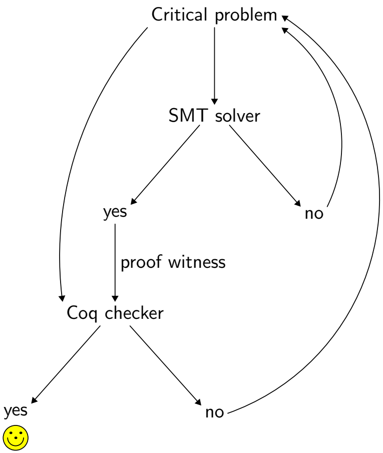
AUTOMATION

SAFETY

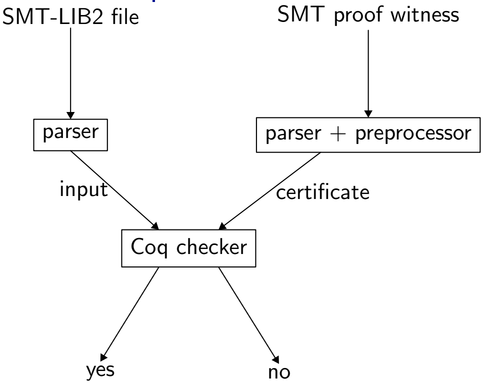
Motivation (2/2)



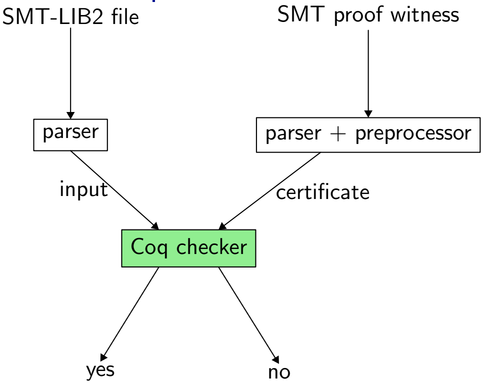
Motivation (2/2)



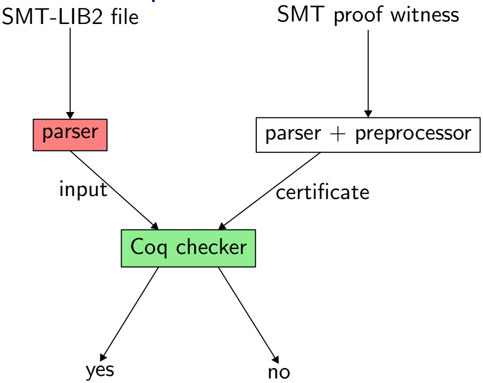
Architecture of SMTCoq



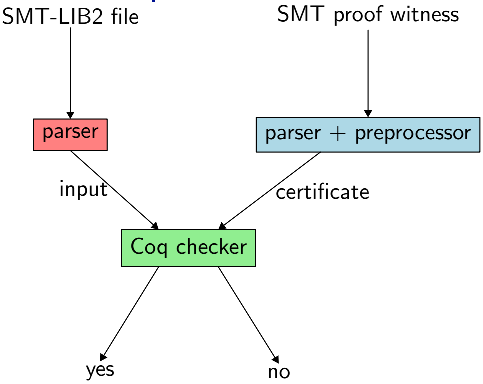
Architecture of SMTCoq



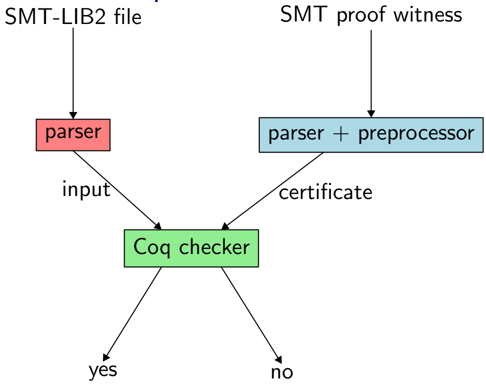
Architecture of SMTCoq



Architecture of SMTCoq



Architecture of SMTCoq



Can be used:

- to certify SMT results
- as Coq tactics
- in larger developments (eg. DP using bit-blasting)

Outline

1 Focus on certificates

2 Focus on the Coq checker

3 Coq tactics

4 Related works

5 Conclusion

SAT case

Decide propositional satisfiability of sets of clauses:

- $x \vee y \qquad x \vee \bar{y} \vee z \qquad \bar{x} \vee z \qquad \bar{z}$

Certificate:

- If satisfiable: assignment of the variables to $\top$ or $\bot$
- If unsatisfiable: proof by resolution of the empty clause

Resolution rule:

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

Examples

Satisfiability of: $x \vee y \qquad x \vee \bar{y} \vee z \qquad \bar{x} \vee z$

$$\{x \mapsto \top, y \mapsto \bot, z \mapsto \top\}$$

Unsatisfiability of: $x \vee y \qquad x \vee \bar{y} \vee z \qquad \bar{x} \vee z \qquad \bar{z}$

$$\dfrac{\dfrac{x \vee y \qquad \dfrac{x \vee \bar{y} \vee z \qquad \bar{z}}{x \vee \bar{y}}}{x} \qquad \dfrac{\bar{x} \vee z \qquad \bar{z}}{\bar{x}}}{\square}$$

Examples

Satisfiability of: $x \vee y \qquad x \vee \bar{y} \vee z \qquad \bar{x} \vee z$

$$\{x \mapsto \top, y \mapsto \bot, z \mapsto \top\}$$

Unsatisfiability of: $x \vee y \qquad x \vee \bar{y} \vee z \qquad \bar{x} \vee z \qquad \bar{z}$

$$\dfrac{\dfrac{x \vee y \qquad \dfrac{x \vee \bar{y} \vee z \qquad \bar{z}}{x \vee \bar{y}}}{x} \qquad \dfrac{\bar{x} \vee z \qquad \bar{z}}{\bar{x}}}{\square}$$

Examples

Satisfiability of: $x \lor y \qquad x \lor \bar{y} \lor z \qquad \bar{x} \lor z$

$$\{x \mapsto \top, y \mapsto \bot, z \mapsto \top\}$$

Unsatisfiability of: $x \lor y \qquad x \lor \bar{y} \lor z \qquad \bar{x} \lor z \qquad \bar{z}$

$$\dfrac{\dfrac{x \lor y \qquad \dfrac{x \lor \bar{y} \lor z \qquad \bar{z}}{x \lor \bar{y}}}{x} \qquad \dfrac{\bar{x} \lor z \qquad \bar{z}}{\bar{x}}}{\square}$$

Resolution chain

SAT modulo Theories

Atoms are now formulas of some theories:

- congruence closure
- linear arithmetic
- ...

Certificate:

- If satisfiable: assignment of the variables
- If unsatisfiable: proof by resolution of the empty clause in which some leaves are theory lemmas

Examples

Satisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z))$

$$\{x \mapsto f(a), y \mapsto a, z \mapsto a\}$$

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{f(x) = f(y)} \qquad f(x) \neq f(y)}{\square}$$

Examples

Satisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z))$

$$\{x \mapsto f(a), y \mapsto a, z \mapsto a\}$$

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{f(x) = f(y)} \qquad f(x) \neq f(y)}{\square}$$

Examples

Satisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z))$

$$\{x \mapsto f(a), y \mapsto a, z \mapsto a\}$$

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\,\dfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{f(x) = f(y)} \qquad\qquad f(x) \neq f(y)}{\square}$$

Outline

1 Focus on certificates

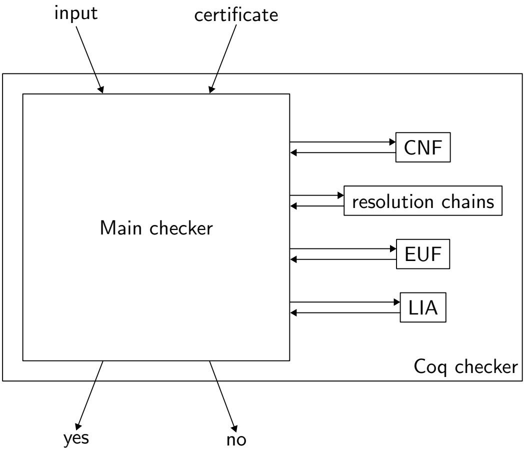
2 Focus on the Coq checker

3 Coq tactics

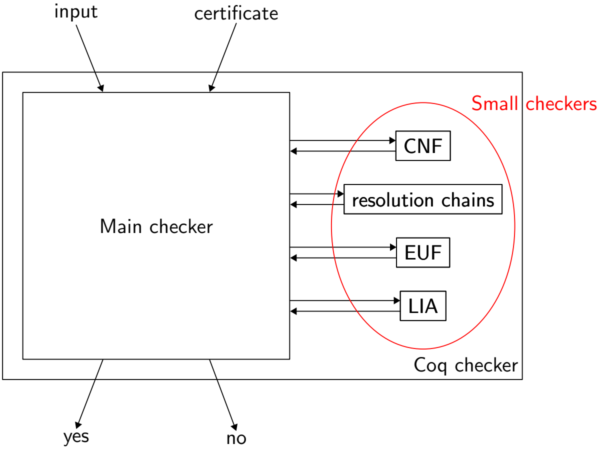
4 Related works

5 Conclusion

A modular checker based on computational reflection



A modular checker based on computational reflection



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 an array of clauses
- sequentially shares out each certificate step between the corresponding small checker
- checks that the last obtained clause is the empty clause

The main checker by example

Unsatisfiability of: $f(x) \neq f(y)$ $f(x) = f(f(z))$ $x = y$

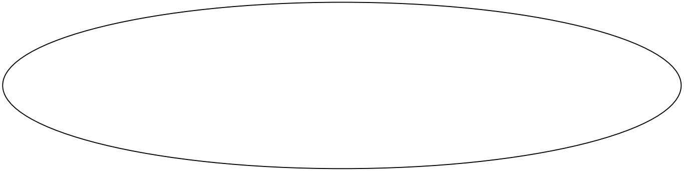
$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{f(x) = f(y)} \qquad\qquad f(x) \neq f(y)}{\square}$$

The main checker by example

Unsatisfiability of: $f(x) \neq f(y)$ $f(x) = f(f(z))$ $x = y$

$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{f(x) = f(y)} \qquad\qquad f(x) \neq f(y)}{\square}$$

A set of clauses:



The main checker by example

Unsatisfiability of: $f(x) \neq f(y)$ $\quad f(x) = f(f(z))$ $\quad x = y$

$$\dfrac{\dfrac{x = y}{} \qquad f(x) \neq f(y)}{}$$

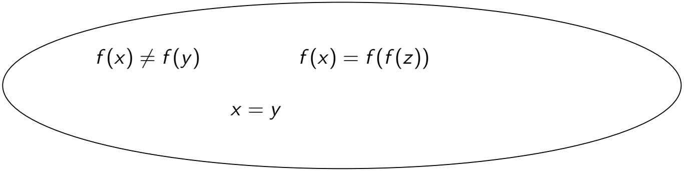
A set of clauses:

The main checker by example

Unsatisfiability of: $f(x) \neq f(y)$ $\quad$ $f(x) = f(f(z))$ $\quad$ $x = y$

$$\frac{\dfrac{x = y}{\qquad\qquad} \qquad f(x) \neq f(y)}{\qquad\qquad\qquad}$$

A set of clauses:



$f(x) \neq f(y)$ $\quad$ $f(x) = f(f(z))$

$x = y$

The main checker by example

Unsatisfiability of: $f(x) \neq f(y)$ $\qquad f(x) = f(f(z))$ $\qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{}$$

A set of clauses:

$f(x) \neq f(y)$ $\qquad f(x) = f(f(z))$

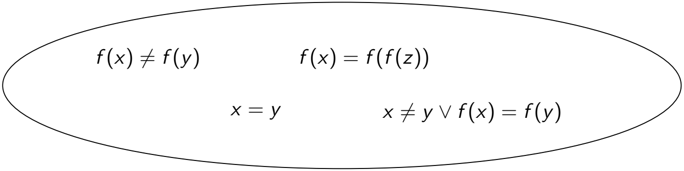
$x = y$

The main checker by example

Unsatisfiability of: $f(x) \neq f(y)$ $f(x) = f(f(z))$ $x = y$

$$\cfrac{\cfrac{\text{EUF}\ \cfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{}$$

A set of clauses:



$f(x) \neq f(y)$ $f(x) = f(f(z))$

$x = y$ $x \neq y \vee f(x) = f(y)$

The main checker by example

Unsatisfiability of: $f(x) \neq f(y)$ $\quad f(x) = f(f(z))$ $\quad x = y$

$$\dfrac{\dfrac{\text{EUF}\,\dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{\square}$$

A set of clauses:

$f(x) \neq f(y)$ $\qquad f(x) = f(f(z))$

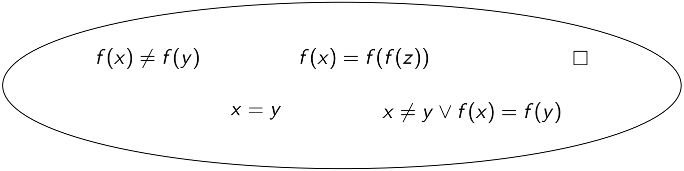
$x = y$ $\qquad x \neq y \vee f(x) = f(y)$

The main checker by example

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{\Box}$$

A set of clauses:



$f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad \Box$

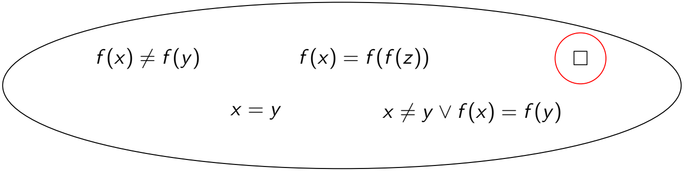
$x = y \qquad x \neq y \lor f(x) = f(y)$

The main checker by example

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{\Box}$$

A set of clauses:



Improvements

Unsatisfiability of: $f(x) \neq f(y)$ $\qquad f(x) = f(f(z))$ $\qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\;\dfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{f(x) = f(y)} \qquad\qquad f(x) \neq f(y)}{\square}$$

Improvements

Unsatisfiability of: $f(x) \neq f(y)$ $\quad f(x) = f(f(z))$ $\quad x = y$

$$\cfrac{\cfrac{\text{EUF}\;\cfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{f(x) = f(y)} \qquad f(x) \neq f(y)}{\square}$$

3 clauses alive at the same time:

Improvements

Unsatisfiability of: $f(x) \neq f(y)$ $f(x) = f(f(z))$ $x = y$

$$\dfrac{\dfrac{x = y}{\qquad} \qquad f(x) \neq f(y)}{\qquad}$$

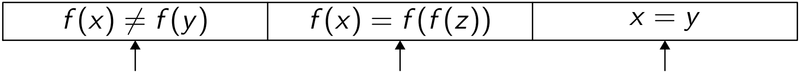
3 clauses alive at the same time:

Improvements

Unsatisfiability of: $f(x) \neq f(y)$ $f(x) = f(f(z))$ $x = y$

$$\dfrac{\dfrac{x = y}{} \qquad f(x) \neq f(y)}{}$$

3 clauses alive at the same time:



$f(x) \neq f(y)$	$f(x) = f(f(z))$	$x = y$

Improvements

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\,\dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{}$$

3 clauses alive at the same time:

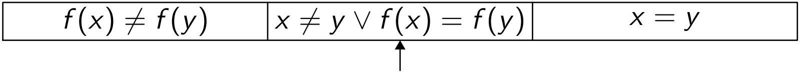
$f(x) \neq f(y)$	$f(x) = f(f(z))$	$x = y$

Improvements

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\,\dfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{}$$

3 clauses alive at the same time:



$f(x) \neq f(y)$	$x \neq y \lor f(x) = f(y)$	$x = y$

Improvements

Unsatisfiability of: $f(x) \neq f(y)$ $f(x) = f(f(z))$ $x = y$

$$\dfrac{\dfrac{\text{EUF}\ \dfrac{}{x \neq y \lor f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{\square}$$

3 clauses alive at the same time:

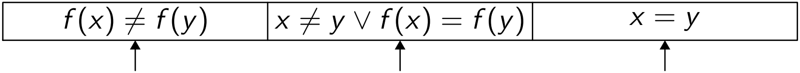
$f(x) \neq f(y)$	$x \neq y \lor f(x) = f(y)$	$x = y$

Improvements

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\mathrm{EUF}\,\dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{\square}$$

3 clauses alive at the same time:



$f(x) \neq f(y)$	$x \neq y \vee f(x) = f(y)$	$x = y$

Improvements

Unsatisfiability of: $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

$$\dfrac{\dfrac{\text{EUF}\;\dfrac{}{x \neq y \vee f(x) = f(y)} \qquad x = y}{} \qquad f(x) \neq f(y)}{\square}$$

3 clauses alive at the same time:

$\square$	$x \neq y \vee f(x) = f(y)$	$x = y$

↑

Small checkers

Current small checkers:

- resolution chains
- CNF computation
- Equality of Uninterpreted Functions
- Linear Integer Arithmetic (using an existing Coq decision procedure)
- Simplifications (eg. $x + 0 \rightsquigarrow x$)

Outline

1 Focus on certificates

2 Focus on the Coq checker

3 Coq tactics

4 Related works

5 Conclusion

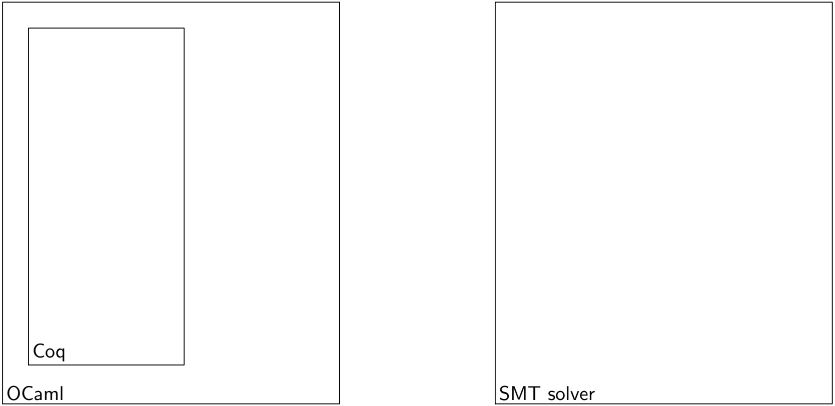
Motivation

Example[1]:

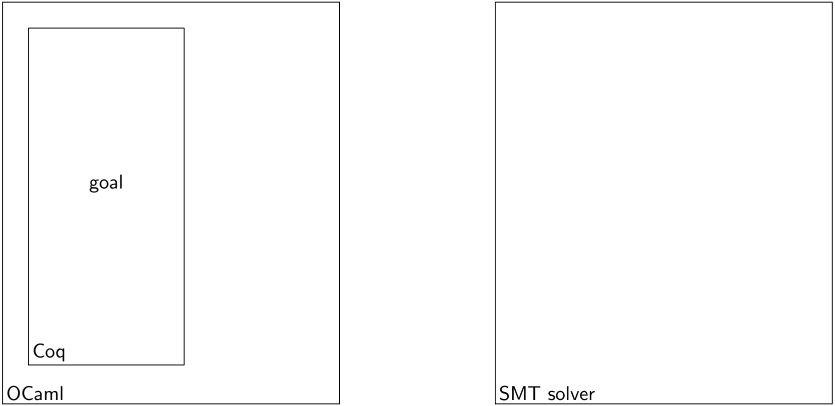
```
Goal forall b1 b2 x1 x2,
  (if b1 then 2 * x1 + 1 else 2 * x1) =
  (if b2 then 2 * x2 + 1 else 2 * x2) ->
    b1 = b2 /\ x1 = x2.
Proof.
  verit.
Qed.
```

[1] Taken from CompCert.

Proof by reflection

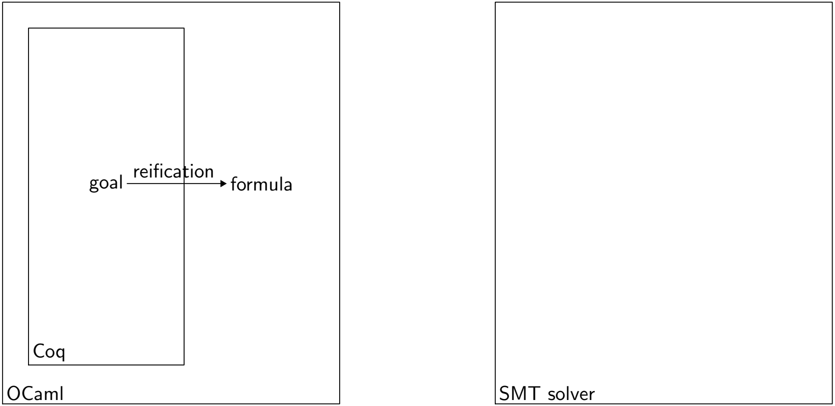


Proof by reflection



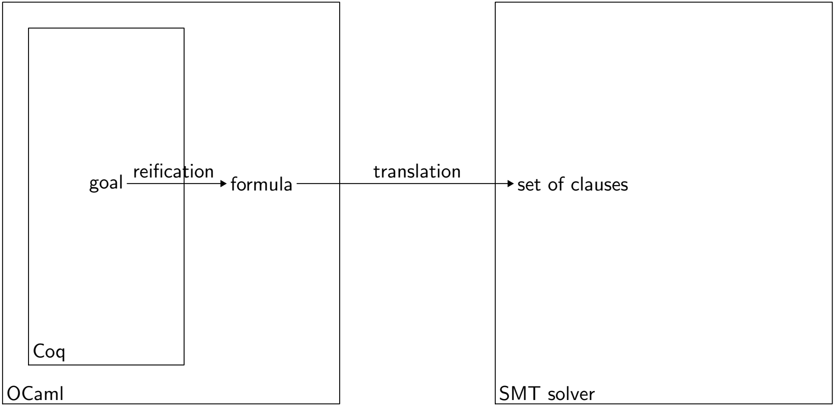
$(\forall \vec{x}, F)$ is true

Proof by reflection



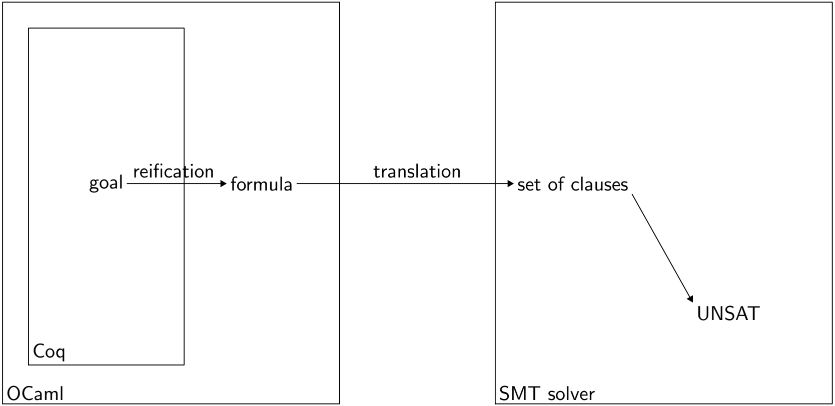
$(\forall \vec{x}, F)$ is true $\Leftrightarrow$ $(\exists \vec{x}, \neg F)$ is false

Proof by reflection



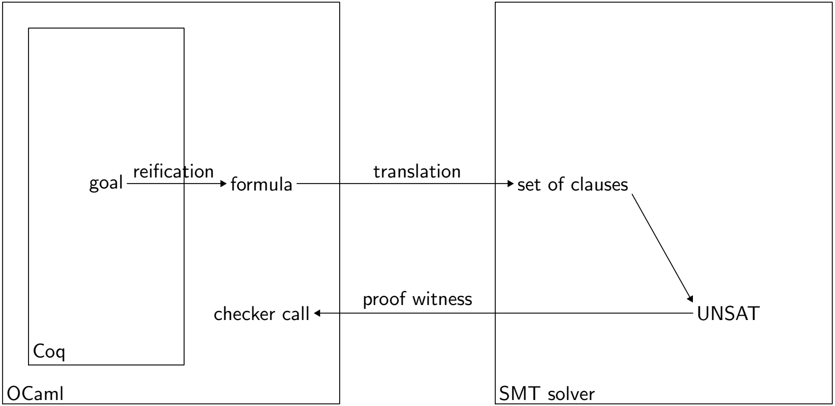
$(\forall \vec{x}, F)$ is true $\Leftrightarrow$ $(\exists \vec{x}, \neg F)$ is false $\Leftrightarrow$ $(\neg F)$ is unsatisfiable

Proof by reflection



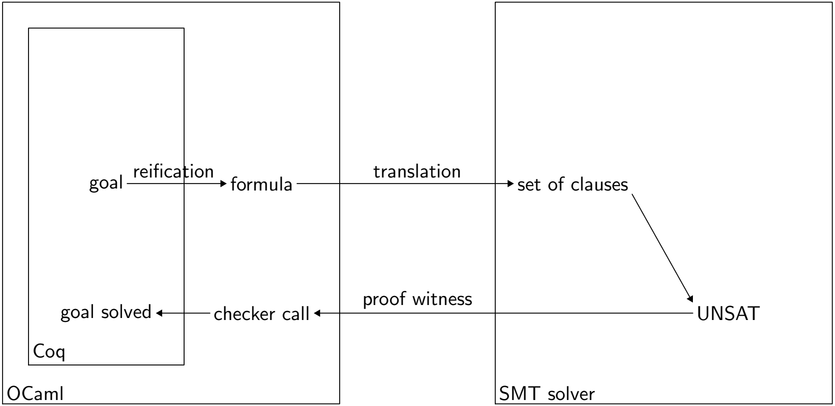
$(\forall\vec{x}, F)$ is true $\Leftrightarrow$ $(\exists\vec{x}, \neg F)$ is false $\Leftrightarrow$ $(\neg F)$ is unsatisfiable

Proof by reflection



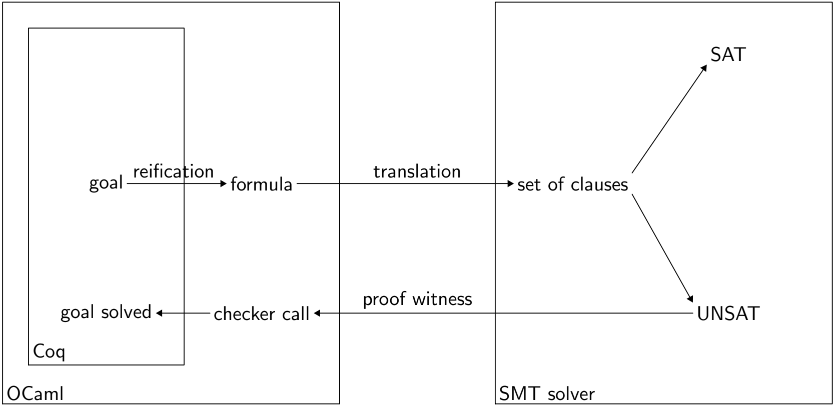
$(\forall \vec{x}, F)$ is true $\Leftrightarrow$ $(\exists \vec{x}, \neg F)$ is false $\Leftrightarrow$ $(\neg F)$ is unsatisfiable

Proof by reflection



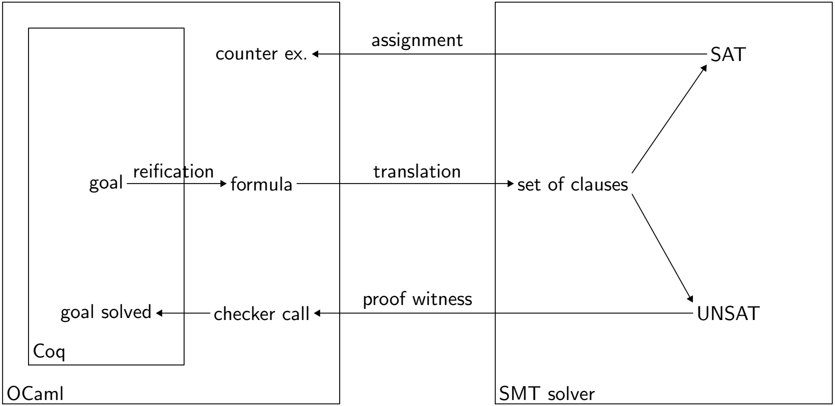
$(\forall \vec{x}, F)$ is true $\Leftrightarrow$ $(\exists \vec{x}, \neg F)$ is false $\Leftrightarrow$ $(\neg F)$ is unsatisfiable

Proof by reflection



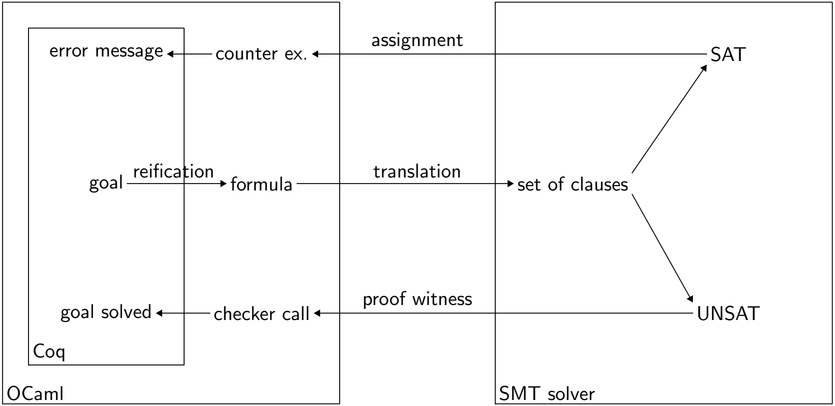
$(\forall \vec{x}, F)$ is true $\Leftrightarrow$ $(\exists \vec{x}, \neg F)$ is false $\Leftrightarrow$ $(\neg F)$ is unsatisfiable

Proof by reflection



$$(\forall \vec{x}, F) \text{ is true} \Leftrightarrow (\exists \vec{x}, \neg F) \text{ is false} \Leftrightarrow (\neg F) \text{ is unsatisfiable}$$

Proof by reflection



$(\forall \vec{x}, F)$ is true $\Leftrightarrow$ $(\exists \vec{x}, \neg F)$ is false $\Leftrightarrow$ $(\neg F)$ is unsatisfiable

What's next

Work in progress

- accept goals in the sort of propositions ($\neq$ Booleans in Coq)
- normalize the goal

Future directions

- handle quantifiers
- encodings before sending to the SMT

Outline

1 Focus on certificates

2 Focus on the Coq checker

3 Coq tactics

4 Related works

5 Conclusion

Another approach

Since Coq is a programming language:

- implement a SMT solver inside
- prove its correctness

↪ followed by S. Lescuyer et al.: embedding Alt-Ergo in Coq (the ergo tactic)

Pros and cons of ergo

Pros:

- a fully certified prover (not *a posteriori*)
- which can be extracted
- self-contained

Cons:

- not robust to small changes
- hard
- likely to be less efficient
- does not benefit from existing tools

Proof reconstruction in Isabelle/HOL

Proof witness verification:

- implemented for zChaff and Z3 in Isabelle/HOL by S. Böhme and T. Weber
- integrated in Sledgehammer by J. Blanchette (currently far more powerful than our tactics)

Pros and cons of Isabelle/HOL

Pros:

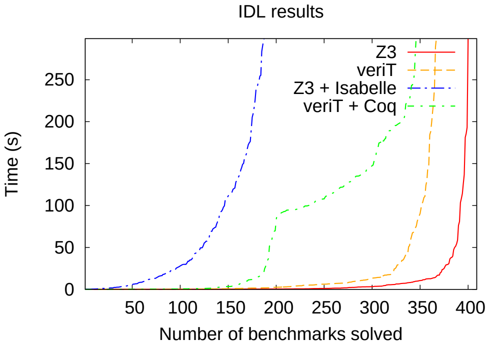
- no proof terms
- smaller trusting base

Cons:

- highly dependent on the format of proof witnesses (here Z3)
- no computational reflection
- no extraction

Benchmarks coming from the SMT-comp

veriT and Z3 on 2000 benchmarks from SMT-LIB



Outline

1 Focus on certificates

2 Focus on the Coq checker

3 Coq tactics

4 Related works

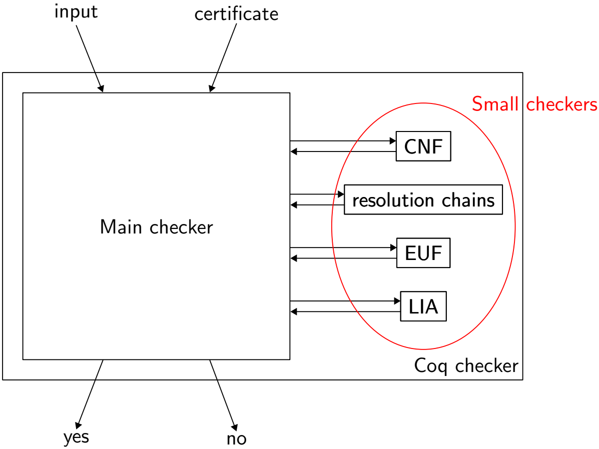
5 Conclusion

Conclusion

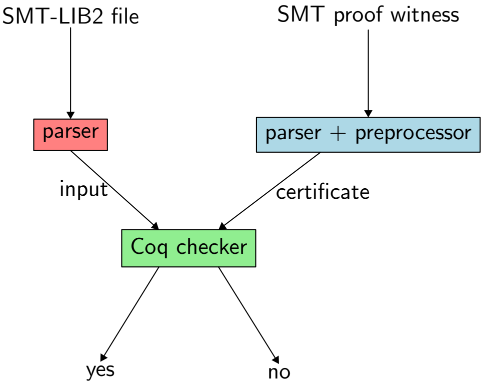
SMTCoq:

- efficient *a posteriori* verification of SMT solvers
  - computational reflection
  - careful choice of term representation
- new decision procedure in Coq
- **modular at many levels**

Small checkers



Integration of new solvers



Advertisement

SMTCoq:

http://www.lix.polytechnique.fr/∼keller/Recherche/smtcoq.html

Certificates:

- our format is a proposal to the SAT/SMT community
- seems like a good balance
- do not hesitate to use it, enhance it...

Perspectives:

- many directions already discusses (new solvers, quantifiers, new theories, encoding of more expressive Coq terms, decision procedure on 31bits integers...)
- confront with applications!