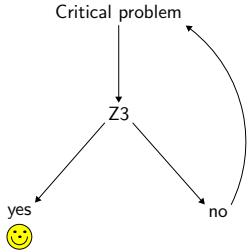
Certification of SMT proof witnesses in the Coq proof assistant

Michaël Armand Benjamin Grégoire Chantal Keller Laurent Théry
Benjamin Werner

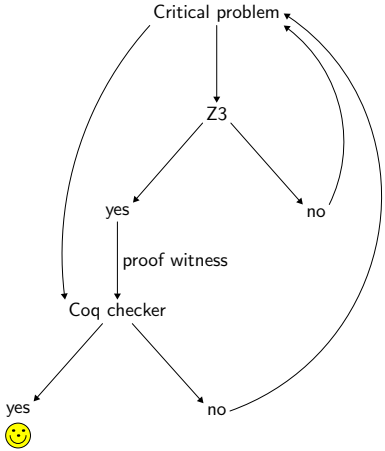
INRIA – École Polytechnique

November, 3rd 2011

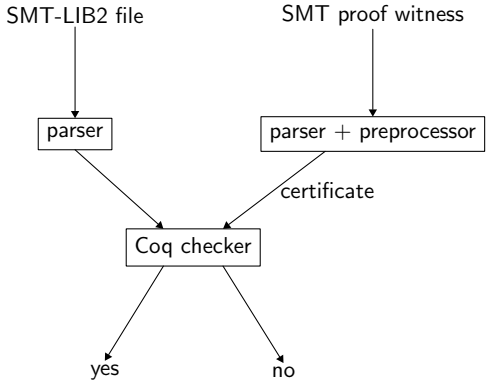
Introduction



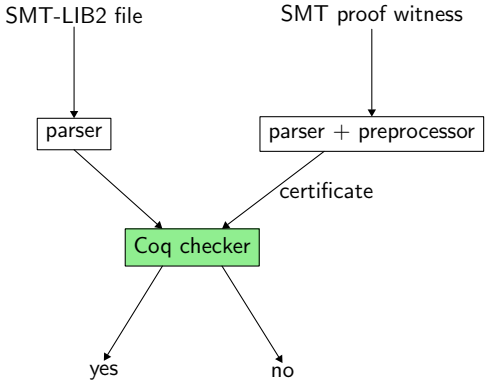
Introduction



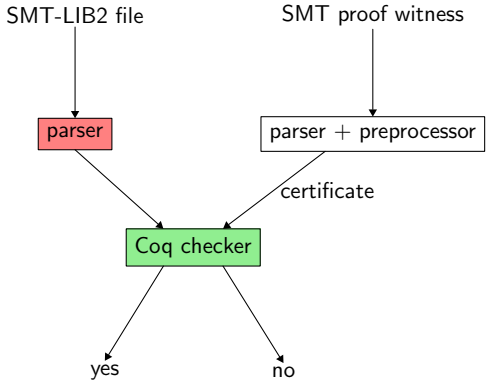
Architecture of SMTCoq



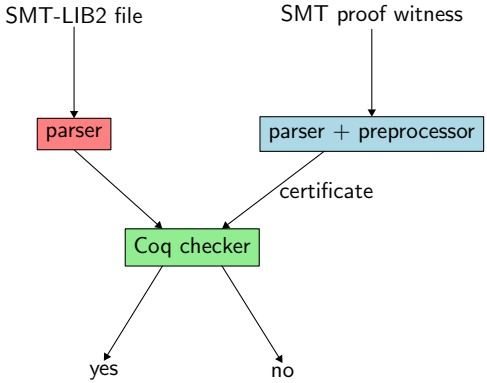
Architecture of SMTCoq



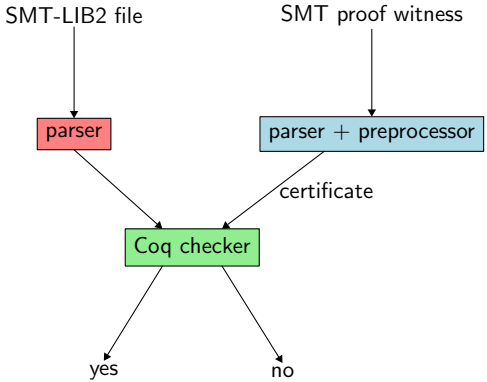
Architecture of SMTCoq



Architecture of SMTCoq



Architecture of SMTCoq



Trusting base: SMT-LIB2 parser + Coq kernel

Outline

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: $x \vee y \quad x \vee \bar{y} \vee z \quad \bar{x} \vee z$

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

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

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

Resolution chain

SAT modulo Theories

Atoms are now formulas of some theories:

- congruence closure
- linear arithmetic
- ...
- CNF computation
- $f(x) \neq f(y) \qquad f(x) = f(f(z)) \qquad x = y$

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

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

Examples

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

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

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

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

Examples

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

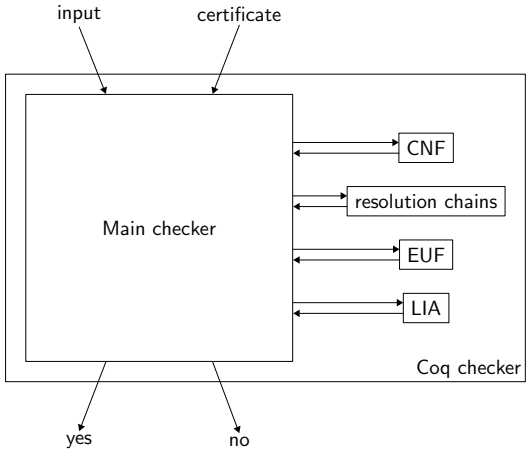
$$\{x \mapsto f(a), y \mapsto a, z \mapsto a\}$$

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

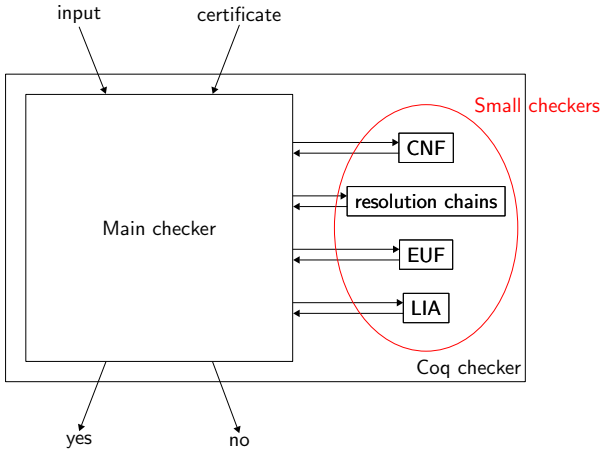
$$\dfrac{\dfrac{x \neq y \vee f(x) = f(y) \qquad x = y}{f(x) = f(y)} \qquad f(x) \neq f(y)}{\square}$$

Outline

A modular checker



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

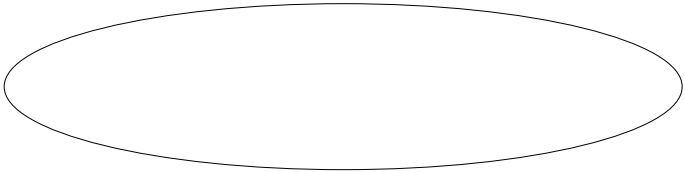
$$\dfrac{\dfrac{x \neq y \vee f(x) = f(y) \qquad x = y}{f(x) = f(y)} \qquad f(x) \neq f(y)}{\square}$$

The main checker by example

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

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

A set of clauses:

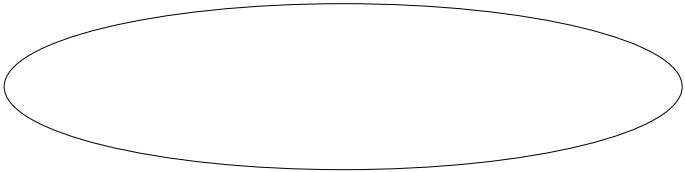


The main checker by example

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

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

A set of clauses:

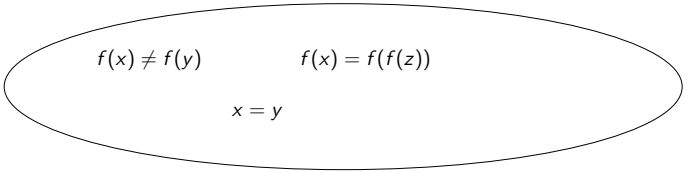


The main checker by example

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

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

A set of clauses:

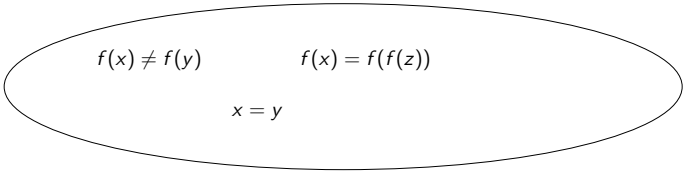


The main checker by example

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

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

A set of clauses:

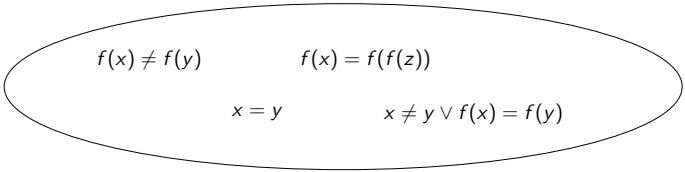


The main checker by example

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

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

A set of clauses:

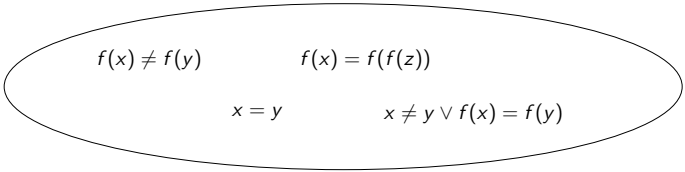


The main checker by example

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

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

A set of clauses:

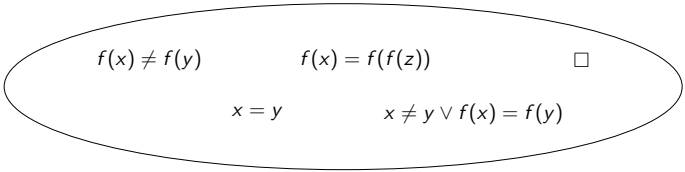


The main checker by example

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

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

A set of clauses:



The main checker by example

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

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

A set of clauses:

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

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

Improvements

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

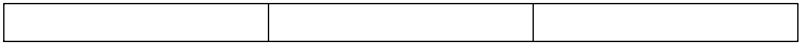
$$\dfrac{\dfrac{x \neq y \vee f(x) = f(y) \qquad x = y}{f(x) = f(y)} \qquad f(x) \neq f(y)}{\square}$$

Improvements

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

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

3 clauses alive at the same time:

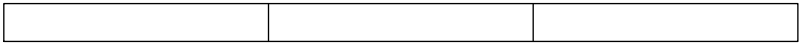


Improvements

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

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

3 clauses alive at the same time:

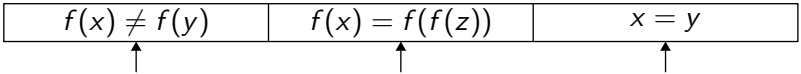


Improvements

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

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

3 clauses alive at the same time:



Improvements

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

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

3 clauses alive at the same time:

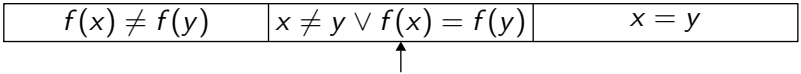
$f(x) \neq f(y)$	$f(x) = f(f(z))$	$x = y$

Improvements

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

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

3 clauses alive at the same time:



Improvements

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

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

3 clauses alive at the same time:

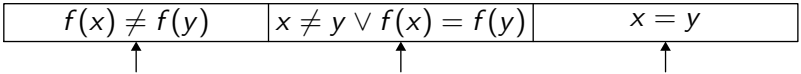
$f(x) \neq f(y)$	$x \neq y \vee f(x) = f(y)$	$x = y$

Improvements

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

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

3 clauses alive at the same time:

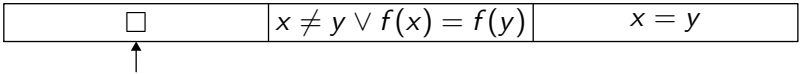


Improvements

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

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

3 clauses alive at the same time:



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

Some related works among others

Proof witness verification:

- S. Böhme and T. Weber: Z3 verification in HOL and Isabelle/HOL
- J. Blanchette: Sledgehammer
- F. Besson et al.: combination of theories in Coq
- P. Fontaine et al.: Harvey in Isabelle/HOL

SMT solvers certification:

- S. Lescuyer et al.: embedding Alt-Ergo in Coq

Differences between Isabelle and Coq

Isabelle:

- based on Deduction
- trusting base: a small set of inference rules
- pro: no proof terms

Coq:

- based on Type Theory
- trusting base: a type checker
- pro: possibility to perform computation (proof by reflection)

Extraction

Coq: a proof assistant and a programming language:

- a programming language: in which the checker is written
- a proof assistant: in which the checher is proved correct

The checker is extracted to OCaml code:

- no need to install Coq to use it
- correctnees remains

Benchmarks coming from the SAT-comp

zChaff on 151 benchmarks from SAT Race'06 and '08

Solved zChaff			Isabelle checker			Coq checker		
#	%	Time	#	%	Time	#	%	Time
79	52	51.9	57	38	100.	79	52	17.5

Benchmarks coming from the SMT-comp

veriT and Z3 on 2000 benchmarks from SMT-LIB

Benchmarks		Solved Z3			Isabelle checker		
Logic	#	#	%	Time	#	%	Time
QF_UF	1852	1834	99	2.5	1775	96	25.8
QF_IDL	409	402	98	0.6	190	46	55.2
QF_LIA	116	107	92	0.7	96	83	46.6

Benchmarks		Solved veriT			Coq checker		
Logic	#	#	%	Time	#	%	Time
QF_UF	1852	1816	98	6.5	1804	97	1.4
QF_IDL	409	368	90	6.3	349	85	37.8
QF_LIA	116	98	84	11.6	98	84	3.1

Proof producing tools

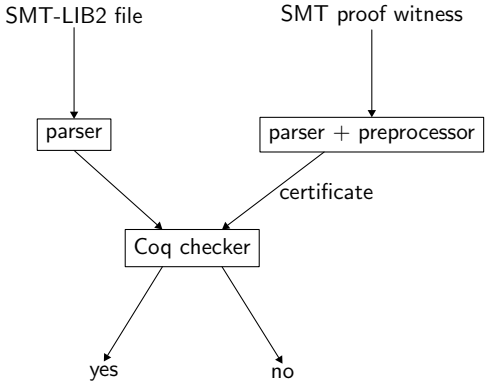
SAT solvers

MiniSat

zChaff

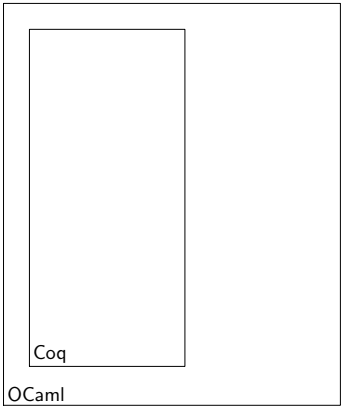
SMT solvers

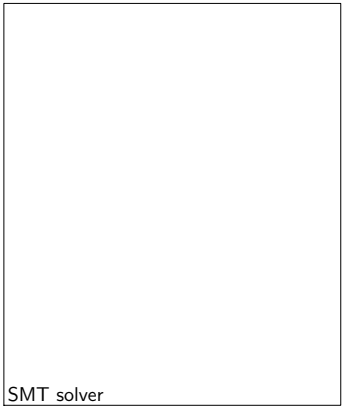
Adding Z3


SMT-LIB2 file
SMT proof witness
parser
parser + preprocessor
certificate
Coq checker
yes
no

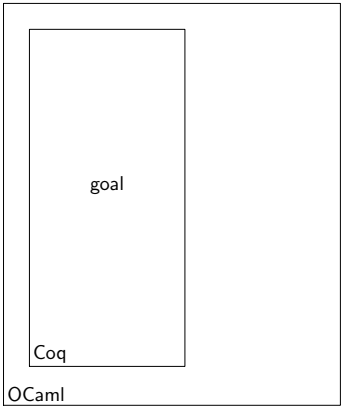
Outline

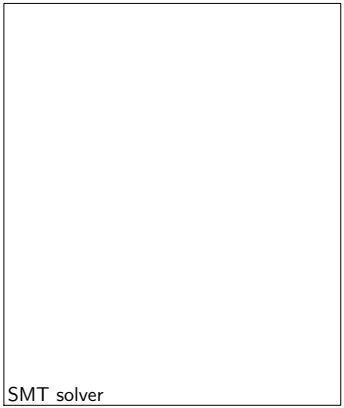
Idea





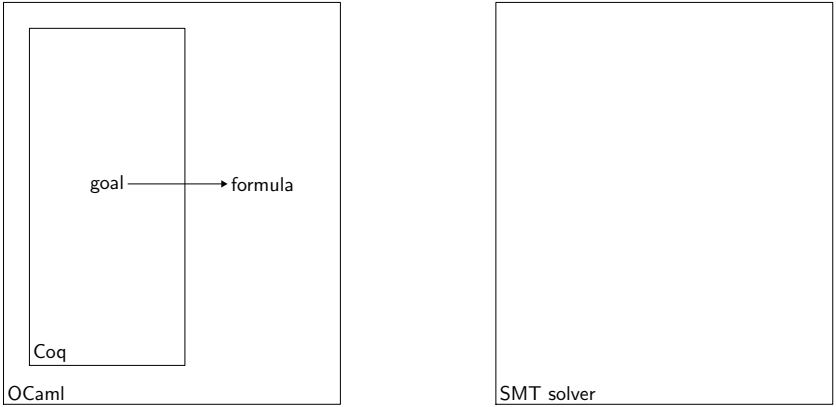
Idea





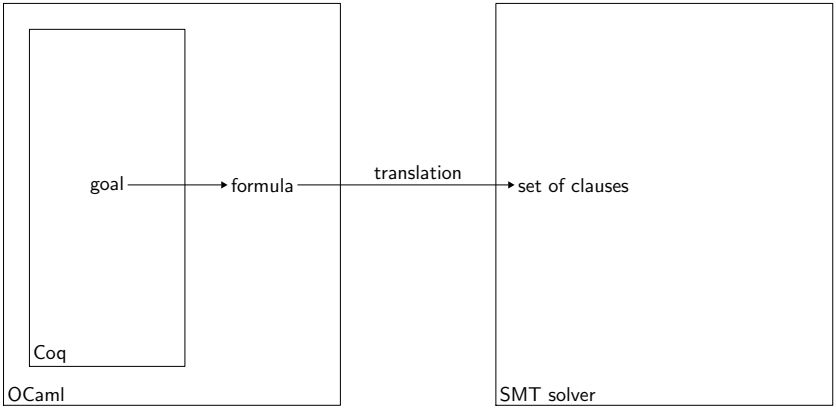
$\forall \vec{x}, F$ is true

Idea



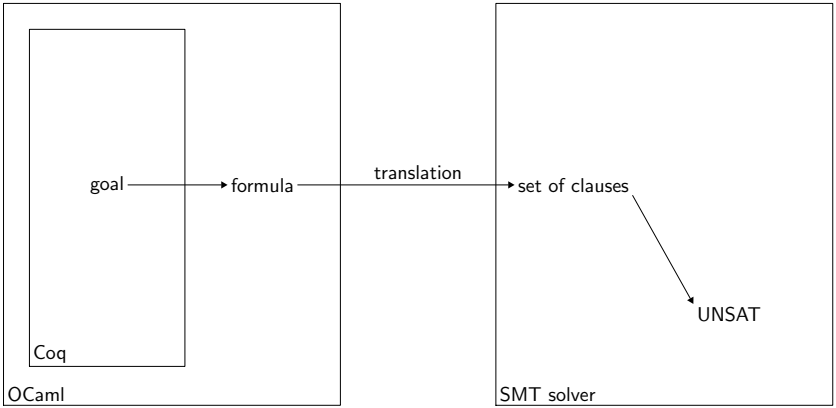
$\forall \vec{x}, F$ is true $\Leftrightarrow \exists \vec{x}, \neg F$ is false

Idea



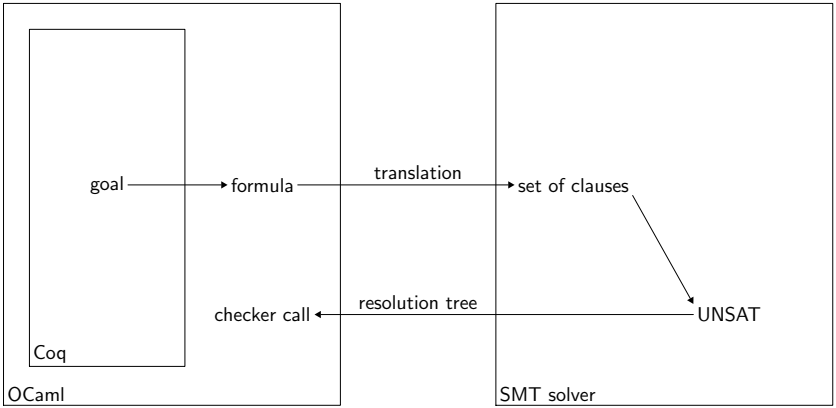
$\forall \vec{x}, F$ is true $\Leftrightarrow \exists \vec{x}, \neg F$ is false $\Leftrightarrow \neg F$ is unsatisfiable

Idea



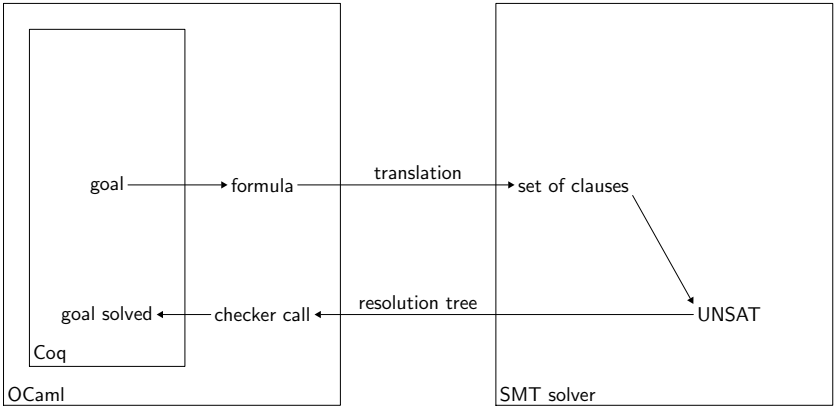
$\forall \vec{x}, F$ is true $\Leftrightarrow \exists \vec{x}, \neg F$ is false $\Leftrightarrow \neg F$ is unsatisfiable

Idea



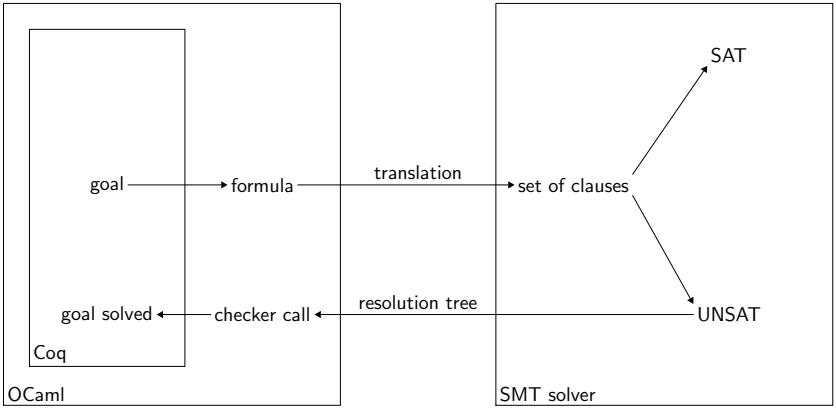
$\forall \vec{x}, F$ is true $\Leftrightarrow \exists \vec{x}, \neg F$ is false $\Leftrightarrow \neg F$ is unsatisfiable

Idea



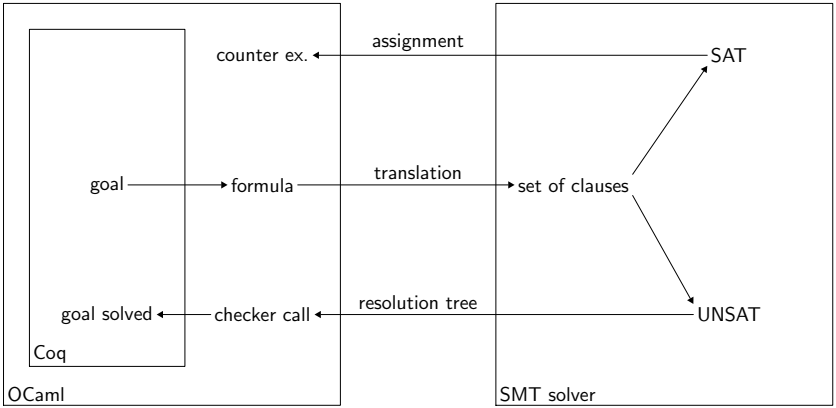
$\forall \vec{x}, F$ is true $\Leftrightarrow \exists \vec{x}, \neg F$ is false $\Leftrightarrow \neg F$ is unsatisfiable

Idea



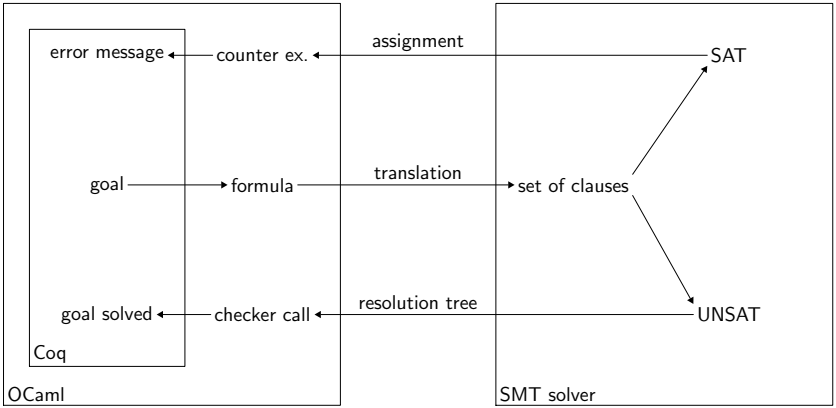
$\forall \vec{x}, F$ is true $\Leftrightarrow \exists \vec{x}, \neg F$ is false $\Leftrightarrow \neg F$ is unsatisfiable

Idea



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

Idea



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

Outline

Conclusion and perspectives

SMTCoq:

- efficient *a posteriori* verification of SMT solvers
- new decision procedure in Coq
- http://www.lix.polytechnique.fr/~keller/Recherche/smtcoq.html

Perspectives:

- Z3
- quantifiers
- new theories
- encoding of more expressive Coq terms