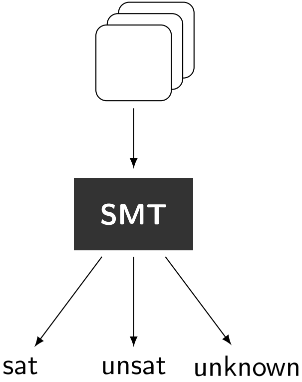
# Proof Reconstruction for Z3 in Isabelle/HOL

Sascha Böhme

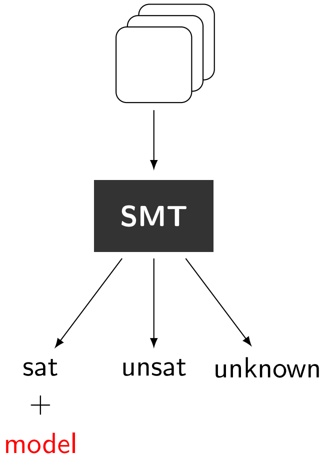
Technische Universität München

3. August 2009




User perspective:

- SMT as “black-box” technology

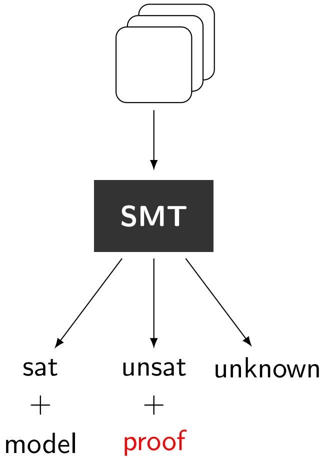



User perspective:

- SMT as "black-box" technology

Additional information:

- satisfiability: model

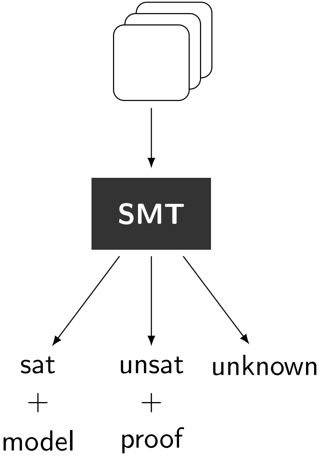



User perspective:

- SMT as “black-box” technology

Additional information:

- satisfiability: model
- unsatisfiability: proof




User perspective:

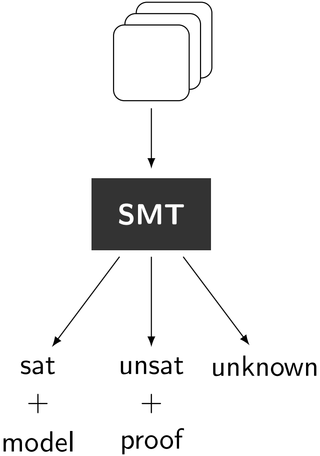
- SMT as “black-box” technology

Additional information:

- satisfiability: model
- unsatisfiability: proof

Increased confidence:

- checkable certificates




User perspective:

- SMT as "black-box" technology

Additional information:

- satisfiability: model
- unsatisfiability: proof

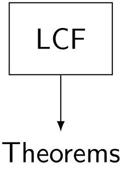
Increased confidence:

- checkable certificates

Our aim:

- certify proofs of Z3
- with Isabelle/HOL

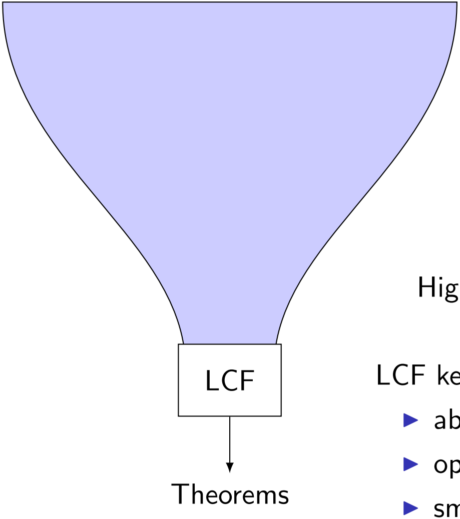
# A Quick Glance at Isabelle/HOL




LCF kernel:

- abstract type: theorems
- operations: basic inference rules
- small

# A Quick Glance at Isabelle/HOL

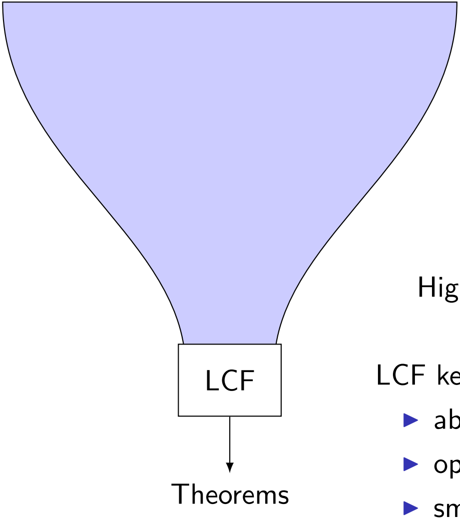



Higher-order logic (HOL)

LCF kernel:

- abstract type: theorems
- operations: basic inference rules
- small

# A Quick Glance at Isabelle/HOL




Proof tools:

- term rewriting (simplifier)
- tableaux prover (blast)
- decision procedures: linear arithmetic, quantifier elimination

Higher-order logic (HOL)

LCF kernel:

- abstract type: theorems
- operations: basic inference rules
- small

# Z3 Terms

Language: many-sorted first-order logic

Terms: $t$, $s$

- variables: $x$, $y$
- applications: $f \ t_1 \ldots t_n$
  - logical connectives: *true*, *false*, $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\sim$
- quantifiers: $\forall$, $\exists$
- terms of sort *bool* (formulas): $P$

# Z3 Terms

Language: many-sorted first-order logic

Terms: $t$, $s$

- variables: $x$, $y$
- applications: $f \; t_1 \ldots t_n$
  - logical connectives: $\mathit{true}$, $\mathit{false}$, $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\sim$
- quantifiers: $\forall$, $\exists$
- terms of sort *bool* (formulas): $P$

Equisatisfiability:

$$(\neg x \vee \mathit{false}) \sim (\neg y) \quad \equiv \quad (\exists x.\, \neg x \vee \mathit{false}) \leftrightarrow (\exists y.\, \neg y)$$

# Z3 Terms

Language: many-sorted first-order logic

Terms: $t$, $s$

- variables: $x$, $y$
- applications: $f\ t_1 \ldots t_n$
  - logical connectives: *true*, *false*, $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\sim$
- quantifiers: $\forall$, $\exists$
- terms of sort *bool* (formulas): $P$

Equisatisfiability:

$$(\neg x \vee \mathit{false}) \sim (\neg y) \quad \equiv \quad (\exists x.\, \neg x \vee \mathit{false}) \leftrightarrow (\exists y.\, \neg y)$$

Natural mapping into higher-order logics (Isabelle/HOL)

- equisatisfiability: representable as equivalence with one exception: Skolemization

# Z3 Proofs

Natural deduction style:

$$\frac{\Gamma_1 \vdash P_1 \qquad \Gamma_2 \vdash P_1 \leftrightarrow P_2}{\Gamma_1 \cup \Gamma_2 \vdash P_2}\ \mathbf{mp}_{\leftrightarrow}$$

# Z3 Proofs

Natural deduction style:

$$\frac{\Gamma_1 \vdash P_1 \qquad \Gamma_2 \vdash P_1 \leftrightarrow P_2}{\Gamma_1 \cup \Gamma_2 \vdash P_2}\ \mathbf{mp}_{\leftrightarrow}$$

Proof trees:

$$\frac{\dfrac{}{\neg \mathit{true} \vdash \neg \mathit{true}}\ \textbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \textbf{rewrite}}{\neg \mathit{true} \vdash \mathit{false}}\ \mathbf{mp}_{\leftrightarrow}$$

# Z3 Proofs

Natural deduction style:

$$\frac{\Gamma_1 \vdash P_1 \qquad \Gamma_2 \vdash P_1 \leftrightarrow P_2}{\Gamma_1 \cup \Gamma_2 \vdash P_2}\ \mathbf{mp}_{\leftrightarrow}$$

Proof trees:

$$\frac{\dfrac{}{\neg \mathit{true} \vdash \neg \mathit{true}}\ \mathbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \mathbf{rewrite}}{\neg \mathit{true} \vdash \mathit{false}}\ \mathbf{mp}_{\leftrightarrow}$$

28 proof rules:

- core logic: **asserted**, **unit**, . . .
- equality: **refl**, **trans**, . . .
- quantifiers: **quant-inst**, **elim-unused**, . . .
- theories: **rewrite**, **th-lemma**

# Proof Reconstruction

$$\dfrac{\dfrac{}{\neg \mathit{true} \vdash \neg \mathit{true}}\ \textbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \textbf{rewrite}}{\neg \mathit{true} \vdash \mathit{false}}\ \textbf{mp}_{\leftrightarrow}$$

# Proof Reconstruction

$$\dfrac{\dfrac{}{\neg \mathit{true} \vdash \neg \mathit{true}}\ \textbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \textbf{rewrite}}{\neg \mathit{true} \vdash \mathit{false}}\ \textbf{mp}_{\leftrightarrow}$$

- bottom-up
- one method for every rule

# Proof Reconstruction

$$\dfrac{\dfrac{}{\neg true \vdash \neg true}\ \textbf{asserted} \qquad \boxed{\dfrac{}{\vdash \neg true \leftrightarrow false}\ \textbf{rewrite}}}{\neg true \vdash false}\ \textbf{mp}_{\leftrightarrow}$$

- bottom-up
- one method for every rule

# Proof Reconstruction

$$\dfrac{\dfrac{}{\neg \mathit{true} \vdash \neg \mathit{true}}\ \textbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \textbf{rewrite}}{\neg \mathit{true} \vdash \mathit{false}}\ \textbf{mp}_{\leftrightarrow}$$

- bottom-up
- one method for every rule

# Proof Reconstruction

$$\dfrac{\dfrac{}{\neg \mathit{true} \vdash \neg \mathit{true}}\ \textbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \textbf{rewrite}}{\neg \mathit{true} \vdash \mathit{false}}\ \textbf{mp}_{\leftrightarrow}$$

- bottom-up
- one method for every rule
- all inferences certified by LCF kernel

# Proof Reconstruction

$$\dfrac{\dfrac{}{\neg \mathit{true} \vdash \boxed{\neg \mathit{true}}}\ \textbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \textbf{rewrite}}{\boxed{\neg \mathit{true}} \vdash \mathit{false}}\ \textbf{mp}_{\leftrightarrow}$$

- bottom-up
- one method for every rule
- all inferences certified by LCF kernel
- additional checks

# Proof Reconstruction

$$\dfrac{\dfrac{}{\neg \mathit{true} \vdash \neg \mathit{true}}\ \textbf{asserted} \qquad \dfrac{}{\vdash \neg \mathit{true} \leftrightarrow \mathit{false}}\ \textbf{rewrite}}{\neg \mathit{true} \vdash \mathit{false}}\ \textbf{mp}_{\leftrightarrow}$$

- bottom-up
- one method for every rule
- all inferences certified by LCF kernel
- additional checks (for debugging)

# Reconstruction Methods

# Reconstruction Methods

- basic inference rules of Isabelle (2 rules)

# Reconstruction Methods

- basic inference rules of Isabelle (2 rules)
- Isabelle theorem and resolution (8 rules)

$$\frac{\Gamma_1 \vdash P_1 \qquad \Gamma_2 \vdash P_1 \leftrightarrow P_2}{\Gamma_1 \cup \Gamma_2 \vdash P_2}\ \mathbf{mp}_{\leftrightarrow}$$

$$P_1 \Longrightarrow P_1 \leftrightarrow P_2 \Longrightarrow P_2$$

# Reconstruction Methods

- basic inference rules of Isabelle (2 rules)
- Isabelle theorem and resolution (8 rules)

$$\frac{\Gamma_1 \vdash P_1 \qquad \Gamma_2 \vdash P_1 \leftrightarrow P_2}{\Gamma_1 \cup \Gamma_2 \vdash P_2}\ \mathbf{mp}_{\leftrightarrow}$$

$$P_1 \Longrightarrow P_1 \leftrightarrow P_2 \Longrightarrow P_2$$

- Isabelle proof tools (simplifier, blast) (9 rules)

# Reconstruction Methods

- basic inference rules of Isabelle (2 rules)
- Isabelle theorem and resolution (8 rules)

$$\frac{\Gamma_1 \vdash P_1 \qquad \Gamma_2 \vdash P_1 \leftrightarrow P_2}{\Gamma_1 \cup \Gamma_2 \vdash P_2}\ \mathbf{mp}_{\leftrightarrow}$$

$$P_1 \Longrightarrow P_1 \leftrightarrow P_2 \Longrightarrow P_2$$

- Isabelle proof tools (simplifier, blast) (9 rules)
- specialized treatment (9 rules)
  - in some cases: optimizations

# Congruence

$$\frac{\Gamma_1 \vdash t_1 = s_1 \qquad \dots \qquad \Gamma_n \vdash t_n = s_n}{\bigcup_{i \leq n} \Gamma_i \vdash f\ t_1 \dots t_n = f\ s_1 \dots s_n}\ \textbf{mono}$$

In principle: provable by simplifier (term rewriting)

But: one of the central rules!

- optimization is worthwhile

Thus: combination of

- congruence: $f = g \Longrightarrow x = y \Longrightarrow f\ x = g\ y$
- reflexivity: $t = t$

# Skolemization

Example:

$$\vdash (\exists x.\, P\ x\ y) \sim P\ (f\ y)\ y$$

# Skolemization

Example:

$$\vdash (\exists x.\, P\ x\ y) \sim P\ (f\ y)\ y$$

With Hilbert choice operator $\varepsilon$:

$$f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash (\exists x.\, P\ x\ y) \leftrightarrow P\ (f\ y)\ y$$

# Skolemization

Example:

$$\vdash (\exists x.\, P\ x\ y) \sim P\ (f\ y)\ y$$

With Hilbert choice operator $\varepsilon$:

$$f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash (\exists x.\, P\ x\ y) \leftrightarrow P\ (f\ y)\ y$$

At the end of reconstruction:

$$\Gamma, f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash \mathit{false}$$

# Skolemization

Example:

$$\vdash (\exists x.\, P\ x\ y) \sim P\ (f\ y)\ y$$

With Hilbert choice operator $\varepsilon$:

$$f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash (\exists x.\, P\ x\ y) \leftrightarrow P\ (f\ y)\ y$$

At the end of reconstruction:

$$\frac{\Gamma, f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash \mathit{false}}{\Gamma \vdash f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \rightarrow \mathit{false}}$$

# Skolemization

Example:

$$\vdash (\exists x.\, P\ x\ y) \sim P\ (f\ y)\ y$$

With Hilbert choice operator $\varepsilon$:

$$f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash (\exists x.\, P\ x\ y) \leftrightarrow P\ (f\ y)\ y$$

At the end of reconstruction:

$$\dfrac{\dfrac{\Gamma, f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash \mathit{false}}{\Gamma \vdash f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \longrightarrow \mathit{false}}}{\Gamma \vdash (\lambda y.\, \varepsilon x.\, P\ x\ y) = (\lambda y.\, \varepsilon x.\, P\ x\ y) \longrightarrow \mathit{false}}$$

# Skolemization

Example:

$$\vdash (\exists x.\, P\ x\ y) \sim P\ (f\ y)\ y$$

With Hilbert choice operator $\varepsilon$:

$$f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash (\exists x.\, P\ x\ y) \leftrightarrow P\ (f\ y)\ y$$

At the end of reconstruction:

$$\cfrac{\cfrac{\cfrac{\Gamma, f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \vdash \mathit{false}}{\Gamma \vdash f = (\lambda y.\, \varepsilon x.\, P\ x\ y) \longrightarrow \mathit{false}}}{\Gamma \vdash (\lambda y.\, \varepsilon x.\, P\ x\ y) = (\lambda y.\, \varepsilon x.\, P\ x\ y) \longrightarrow \mathit{false}}}{\Gamma \vdash \mathit{false}}$$

# Theories

Rewriting (**rewrite**):

$$\frac{}{\vdash\ f\ \ t_1 \ldots t_n = s}$$

- in general: apply rules of $f$
- simplifier, linear arithmetic, specialized procedures

# Theories

Rewriting (**rewrite**):

$$\frac{}{\vdash \ f \ t_1 \dots t_n = s}$$

- in general: apply rules of $f$
- simplifier, linear arithmetic, specialized procedures

Theory reasoning (**th-lemma**):

$$\frac{}{\vdash \ \bigvee_{i \in I} P_i} \qquad \frac{\Gamma_1 \vdash P_1 \qquad \dots \qquad \Gamma_n \vdash P_n}{\bigcup_{i \leq n} \Gamma_i \vdash \mathit{false}}$$

- linear arithmetics: Fourier-Motzkin elimination
- arrays: simplifier

# Experimental Results

- 5 SMT-LIB logics
- 100 unsatisifiably benchmarks (randomly selected)
- timeout: Z3: 2 minutes, Isabelle/HOL: 10 minutes

| Logic | Solved by Z3 | Reconstruction | | | Factor |
|---|---|---|---|---|---|
| | | Success | Failure | Timeout | |
| QF_UF | 96 | 33 | 27 | 36 | 6.5 |
| QF_UFLIA | 99 | 93 | 0 | 6 | 29.6 |
| QF_UFLRA | 100 | 43 | 0 | 57 | 558.3 |
| AUFLIA | 100 | 50 | 31 | 19 | 81.3 |
| AUFLIRA | 100 | 81 | 6 | 13 | 24.3 |

Bad performance:

- only few optimizations implemented
- huge formulas of benchmarks

# Proof Reconstruction Failures

- Incomplete documentation of **rewrite**:

  *The head function symbol of the left-hand side is interpreted.*

# Proof Reconstruction Failures

- Incomplete documentation of **rewrite**:

  *The head function symbol of the left-hand side is interpreted.*

  But:

$$\frac{}{\vdash\ P \wedge (\forall x : \mathit{int}.\, x > 0)\ \leftrightarrow\ \mathit{false} \wedge P}\ \textbf{rewrite}$$

# Proof Reconstruction Failures

- Incomplete documentation of **rewrite**:
  *The head function symbol of the left-hand side is interpreted.*
  But:

$$\dfrac{}{\vdash\ P \wedge (\forall x : \mathit{int}.\, x > 0) \ \leftrightarrow\ \mathit{false} \wedge P}\ \textbf{rewrite}$$

$$\dfrac{}{\vdash\ (P_1 \wedge P_2) \ \leftrightarrow\ \neg(\neg P_1 \vee \neg P_2)}\ \textbf{rewrite}$$

# Proof Reconstruction Failures

- Incomplete documentation of **rewrite**:
  *The head function symbol of the left-hand side is interpreted.*
  But:

$$\frac{}{\vdash\ P \wedge (\forall x : \mathit{int}.\, x > 0)\ \leftrightarrow\ \mathit{false} \wedge P}\ \textbf{rewrite}$$

$$\frac{}{\vdash\ (P_1 \wedge P_2)\ \leftrightarrow\ \neg(\neg P_1 \vee \neg P_2)}\ \textbf{rewrite}$$

- Unit resolution:

$$\frac{\Gamma_1 \vdash\ P_1 \vee P_2 \vee P_1 \qquad \Gamma_2 \vdash\ \neg P_2}{\Gamma_1 \cup \Gamma_2 \vdash\ P_1}\ \textbf{unit}$$

# Proof Reconstruction Failures

- Incomplete documentation of **rewrite**:

  *The head function symbol of the left-hand side is interpreted.*

  But:

$$\frac{}{\vdash\ P \wedge (\forall x : \mathit{int}.\, x > 0)\ \leftrightarrow\ \mathit{false} \wedge P}\ \textbf{rewrite}$$

$$\frac{}{\vdash\ (P_1 \wedge P_2)\ \leftrightarrow\ \neg(\neg P_1 \vee \neg P_2)}\ \textbf{rewrite}$$

- Unit resolution:

$$\frac{\Gamma_1 \vdash\ P_1 \vee P_2 \vee P_1 \qquad \Gamma_2 \vdash\ \neg P_2}{\Gamma_1 \cup \Gamma_2 \vdash\ P_1}\ \textbf{unit}$$

- Transitivity:

$$\frac{\Gamma_1 \vdash\ s = t \qquad \Gamma_2 \vdash\ u = t}{\Gamma_1 \cup \Gamma_2 \vdash\ s = u}\ \textbf{trans}$$

# Conclusion

Proof reconstruction for Z3:

- in Isabelle/HOL: certification by LCF kernel
- challenges: equisatisfiability, huge formulas
- helped to debug Z3 proof generation

Future work:

- improve performance
- integrate into Isabelle/HOL
- consider further theories