max planck institut informatik

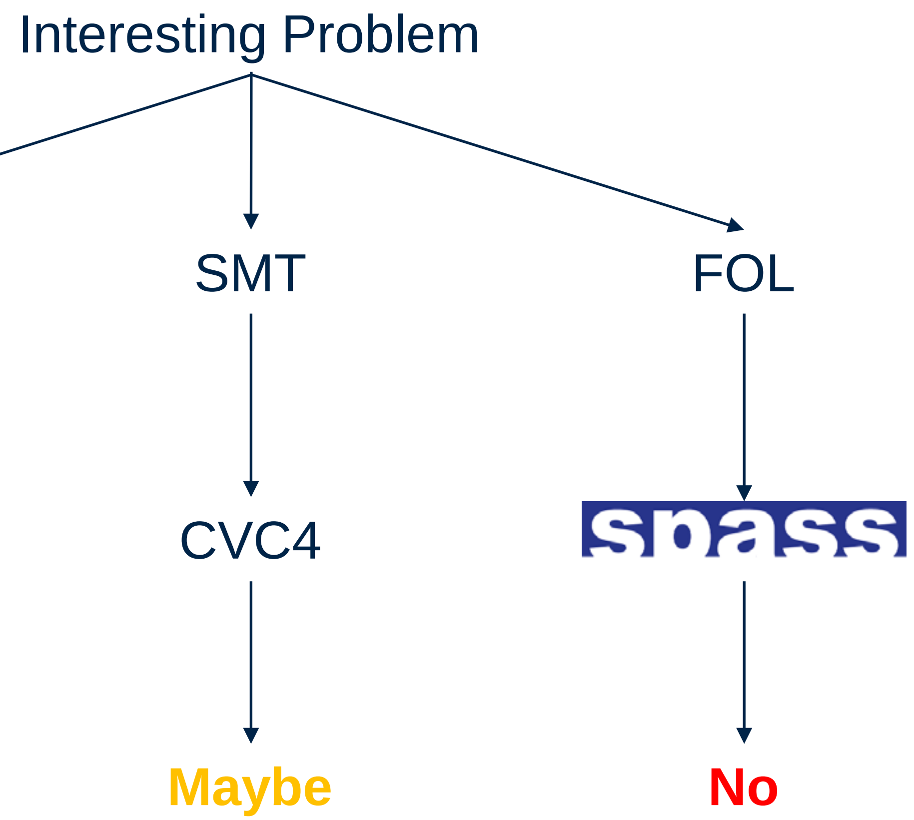
State-of-the-art

FOL

Solving

Christoph Weidenbach

Case Study



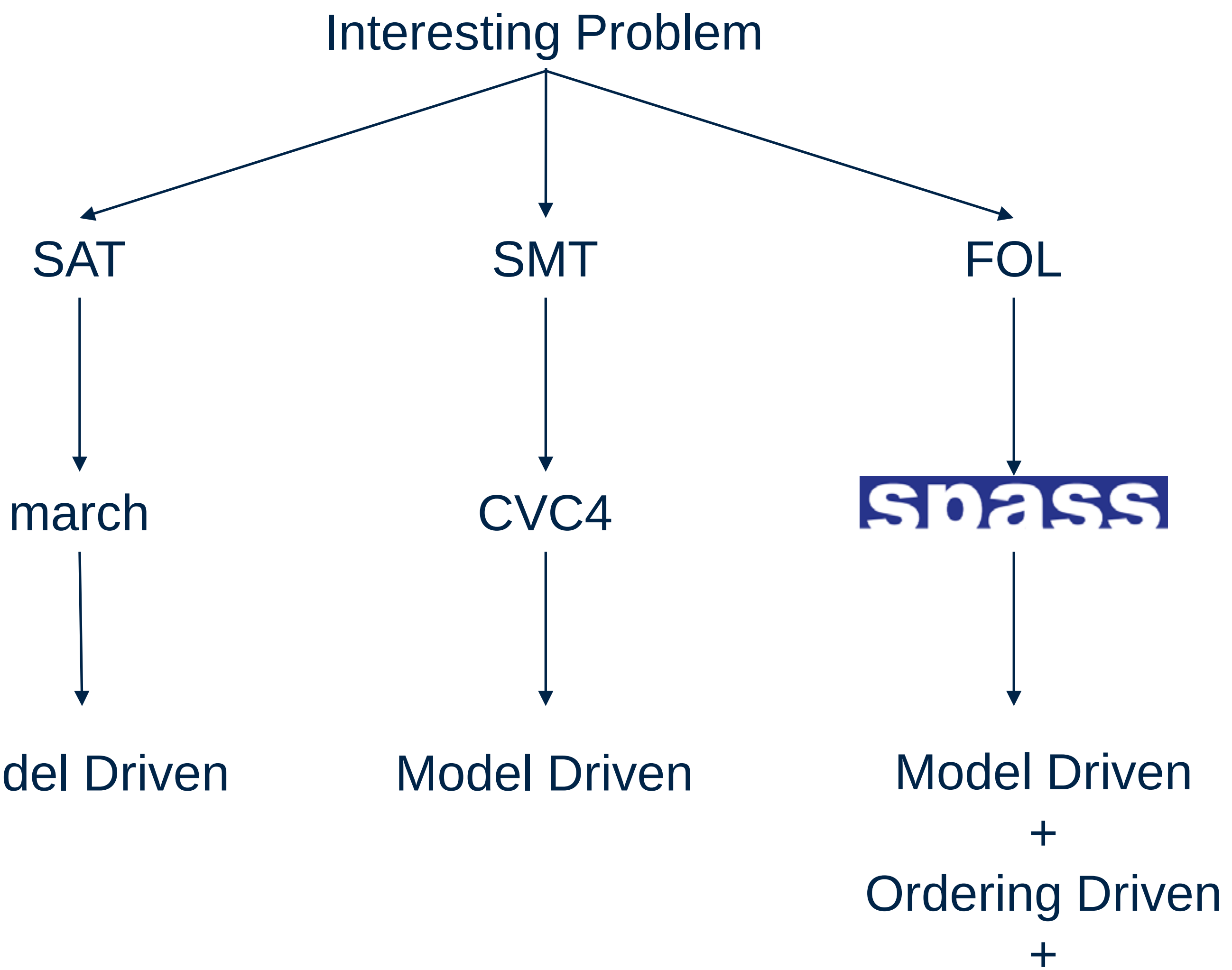
Naive | Some Knowledge | Expert

max planck institut
informatik

Why State-of-the-art FOL Solving does not work in general

Christoph Weidenbach

Reasoning


Interesting Problem
SAT
SMT
FOL
march
CVC4
spass
Model Driven
Model Driven
Model Driven
+
Ordering Driven
+

FOL PCP

$$\neg R(f(x), f(x))$$
$$\neg R(g(x), g(x))$$
$$R(\epsilon, \epsilon)$$
$$\neg R(x, y) \vee R(f(f(x)), f(y))$$
$$\neg R(x, y) \vee R(g(x), f(f(g(y))))$$

PCP (Emil Post 192x:1946)

1	2
ff	g
f	ffg

$$2, 1, 1$$

$$ff\ ff\ g$$
$$f\ f\ ffg$$

$$R(\epsilon, \epsilon)$$

$$R(g(\epsilon), f(f(g(\epsilon))))$$

$$R(f(f(g(\epsilon))), f(f(f(g(\epsilon)))))$$

$$R(f(f(f(f(g(\epsilon))))), f(f(f(f(g(\epsilon))))))$$

$$\bot$$

SMT PCP

$$\neg R(f(x), f(x))$$
$$\neg R(g(x), g(x))$$
$$R(\epsilon, \epsilon)$$
$$\neg R(x, y) \lor R(f(f(x)), f(y))$$
$$\neg R(x, y) \lor R(g(x), f(f(g(y))))$$

$$R(\epsilon, \epsilon)$$
$$R(g(\epsilon), f(f(g(\epsilon))))$$
$$R(f(f(g(\epsilon))), f(f(f(g(\epsilon)))))$$

$$R(f(f(f(f(g(\epsilon))))), f(f(f(f(g(\epsilon))))))$$

$$\bot$$

SMT Model

$$\neg R(f(\mathtt{a}), f(\mathtt{a}))$$
$$\neg R(g(\mathtt{b}), g(\mathtt{b}))$$
$$R(\epsilon, \epsilon)$$
$$\neg R(\mathtt{c}_1, \mathtt{d}_1) \lor R(f(f(\mathtt{c}_1)), f(\mathtt{d}_1))$$
$$\neg R(\mathtt{c}_2, \mathtt{d}_2) \lor R(g(\mathtt{c}_2), f(f(g(\mathtt{d}_2))))$$

$$R(\epsilon, \epsilon)$$
$$\{\mathtt{c}_2 \mapsto \epsilon,\ \mathtt{d}_2 \mapsto \epsilon\}$$
$$\{\mathtt{c}_1 \mapsto g(\epsilon),\ \mathtt{d}_1 \mapsto f(f(g(\epsilon)))\}$$

STOP

SAT PCP

$\neg R(f(a), f(a))$	P_1
$\neg R(g(b), g(b))$	P_2
$R(\epsilon, \epsilon)$	Q
$\neg R(g(\epsilon), f(f(g(\epsilon)))) \vee R(f(f(g(\epsilon))), f(f(f(g(\epsilon)))))$	$\neg S \vee T$
$\neg R(\epsilon, \epsilon) \vee R(g(\epsilon), f(f(g(\epsilon))))$	$\neg Q \vee S$

SAT Model: $[Q, S, T, P_1, P_2]$

FOL PCP Again

$$\neg R(f(x), f(x))$$
$$\neg R(g(x), g(x))$$
$$R(\epsilon, \epsilon)$$
$$\neg R(x, y) \lor R(f(f(x)), f(y))$$
$$\neg R(x, y) \lor R(g(x), f(f(y)))$$

Superposition Model Operator

$$R(\epsilon, \epsilon)$$
$$\neg R(f(\epsilon), f(\epsilon))$$
$$\neg R(g(\epsilon), g(\epsilon))$$
$$R(g(\epsilon), f(f(\epsilon)))$$
$$R(f(f(\epsilon)), f(\epsilon))$$
$$R(f(f(g(\epsilon))), f(f(f(\epsilon))))$$
$$R(g(g(\epsilon)), f(f(f(f(\epsilon)))))$$
$$\vdots$$

Abstract Model Summary

max planck institut informatik

	Model Size	# Models	Computation
FOL	infinite	not enumerable	undecidable
SMT	exponential	exponential	decidable
SAT	linear	exponential	linear

“FOL Model-Based Reasoning is Difficult,”

“and sometimes it is not the Right Approach,”

“and there is no single Right Approach.”

Why FOL Reasoning?

more expressive

$$\neg R(f(x), f(x))$$
$$\neg R(g(x), g(x))$$
$$R(\epsilon, \epsilon)$$
$$\neg R(x, y) \vee R(f(f(x)), g(y))$$
$$\neg R(x, y) \vee R(g(x), f(f(y)))$$

Orderings + Redundancy

Superposition

No Inference between Maximal Literals

There is a Model

$$\neg R(x, y) \vee R(f(f(x)), g(y)) \qquad \neg R(x, y) \vee R(f^{2n}(x), g^{n}(y)) \vee C$$

Redundancy/Simplification $\neg R(x, y) \vee C$

Why FOL Reasoning?

$$R(x,y) \lor R(y,x)$$
$$\neg R(x,y) \lor R(y,x)$$
$$\neg R(a_1,a_2) \lor \neg R(a_2,a_3) \lor \ldots \lor \neg R(a_{n-1},a_n)$$

$R(x,y)$ Proof of $n+1$ Steps

$$R(a_1,a_2) \lor R(a_2,a_1)$$
$$\neg R(a_2,a_1) \lor R(a_1,a_2)$$

$R(a_1,a_2)$ Proof of $2n$ Steps

In general: Exponential Gap (Counter n-ary Predicate)

CDCL Reasoning

$$N = \{\ldots, P \vee \neg Q \vee R,\ \neg S \vee \neg Q \vee \neg R, \ldots\}$$

$$([\ldots S^i, \ldots, \neg P^k], k, \top)$$

$$\Rightarrow_{\text{CDCL}} ([\ldots S^i, \ldots, \neg P^k, Q^{k+1}], k+1, \top)$$

$$\Rightarrow_{\text{CDCL}} ([\ldots S^i, \ldots, \neg P^k, Q^{k+1}, R^{P \vee \neg Q \vee R}], k+1, \top)$$

$$\Rightarrow_{\text{CDCL}} ([\ldots S^i, \ldots, \neg P^k, Q^{k+1}, R^{P \vee \neg Q \vee R}], k+1, \neg S \vee \neg Q \vee \neg R)$$

$$\Rightarrow_{\text{CDCL}} ([\ldots S^i, \ldots, \neg P^k, Q^{k+1}], k+1, \neg S \vee \neg Q \vee P)$$

Learn: $\neg S \vee \neg Q \vee P$

$N^{\prec \neg S \vee \neg Q \vee P} \not\models \neg S \vee \neg Q \vee P$ Not Redundant NP-hard

$\ldots \prec S \prec \ldots \prec P \prec Q \prec R$ Termination

Superposition Propositional

Ordering: $\ldots \prec S \prec \ldots \prec P \prec Q \prec R$

Superposition Left

$(N \uplus \{C_1 \vee P, C_2 \vee \neg P\}) \Rightarrow_{\text{SUP}} (N \cup \{C_1 \vee P, C_2 \vee \neg P\} \cup \{C_1 \vee C_2\})$

where

(i) P is strictly maximal in $C_1 \vee P$

(ii) $\neg P$ is maximal in $C_2 \vee \neg P$

$$P \vee \neg Q \vee R,\ \neg S \vee \neg Q \vee \neg R \Rightarrow_{\text{SUP}} \neg S \vee \neg Q \vee P$$

Redundancy: C redundant if $N^{\prec C} \models C$ or $C \in N$

Superposition FOL

Ordering: $\ldots \prec R(x,y) \prec R(f(x),f(y)) \prec R(f(x),f(f(y))) \prec \ldots$

Superposition Left

$(N \uplus \{C_1 \lor A, C_2 \lor \neg B\})$

$\Rightarrow_{\text{SUP}} (N \cup \{C_1 \lor A, C_2 \lor \neg B\} \cup \{(C_1 \lor C_2)\sigma\})$

where

(i) $A\sigma = B\sigma$

(i) $A\sigma$ is strictly maximal in $(C_1 \lor A)\sigma$

(ii) $\neg B\sigma$ is maximal in $(C_2 \lor \neg B)\sigma$

$\neg R(f(x'), f(f(y'))),\ \neg R(x,y) \lor R(f(x),f(y)) \Rightarrow_{\text{SUP}} \neg R(x', f(y'))$

$\sigma = \{x \mapsto x', y \mapsto f(y')\}$

Superposition Calculus

Completeness:
For some unsatisfiable set of first-order clauses N, Superposition derives $\perp$ after finitely many steps.

Models:
If Superposition terminates on N with respect to redundancy, then N is satisfiable and the Superposition Model Operator generates a model for N.

[BachmairGanzinger]

Superposition PCP

$\neg R(f(x), f(x))$

$\neg R(g(x), g(x))$

$R(\epsilon, \epsilon)$

$\neg R(x', y') \vee R(f(f(x')), f(y'))$

$\neg R(x', y') \vee R(g(x'), f(f(g(y'))))$

$\Rightarrow_{\text{SUP}} \neg R(x, f(x))$ $\quad \sigma = \{x \mapsto f(x'), y' \mapsto f(x')\}$

$\Rightarrow_{\text{SUP}} \neg R(x, f(f(x)))$ $\quad \sigma = \{x \mapsto f(f(x')), y' \mapsto f(f(x'))\}$

$\Rightarrow_{\text{SUP}} \neg R(x, x)$ $\quad \sigma = \{x \mapsto g(x'), y' \mapsto x'\}$

$\Rightarrow_{\text{SUP}} \perp$ $\quad \sigma = \{x \mapsto \epsilon\}$

Superposition PCP

$$\neg R(f(x), f(x))$$
$$\neg R(g(x), g(x))$$
$$R(\epsilon, \epsilon)$$
$$\neg R(x, y) \vee R(f(f(x)), g(y))$$
$$\neg R(x, y) \vee R(g(x), f(f(y)))$$

$$R(\epsilon, \epsilon)$$
$$\neg R(f(\epsilon), f(\epsilon))$$
$$\neg R(g(\epsilon), g(\epsilon))$$
$$R(g(\epsilon), f(f(\epsilon)))$$
$$R(f(f(\epsilon)), g(\epsilon))$$
$$\vdots$$

SAT/SMT cannot determine satisfiability.

SUP Decision Procedure

- Monadic Fragment (with Equality)
- Guarded Fragment (with Equality)
- Two-Variable Fragment (with Equality)
- Maslov's classes $\overline{K}$, $\overline{DK}$
- Fluted Logic
- Bernays-Schoenfinkel ($\exists^*\forall^*$) (with Equality)
- Description Logics
- Modal Logics, Non-Classical Logics (PIDL)
- Shallow (Non-)Linear Monadic Clauses (+ Equations)
- Data Structures (Lists, Arrays, …)

[Bachmair, Bonancina, Hillenbrand, Hustadt, Fermüller, Ganzinger, Georgieva, Jacquemard, Leitsch, Meyer, Motik, Nivelle, Nieuvenhuis, Pratt-Hartmann, Ranise, Rijke, Rubio, Rusinowitch, Sattler, Schmidt, Schulz, Tammet, Teucke, Veanes, Waldmann, Weidenbach]

Bernays-Schoenfinkel (BS)

Finitely many constants: $a_1, a_2, \ldots, a_n$ n fixed

No non-constant function symbols

$\neg R(x,y) \lor \neg R(y,z) \lor R(x,z)$

$\neg R(x,x)$

NEXPTIME-Complete

$\mathrm{Ancestor}(\mathrm{Paul}, \mathrm{John})$

$\neg\,\mathrm{Father}(x,y) \lor \mathrm{Ancestor}(x,y)$

Target language for a number of decidable logics: QBF, Description Logics, Modal Logics, Monadic Fragment, Ontology Languages

“and there is no single Right Approach.”

BS to SAT one Constant

Finitely many constants: a_1

$\neg R(a_1, a_1) \lor \neg R(a_1, a_1) \lor R(a_1, a_1)$ $\neg R \lor \neg R \lor R$

$\neg R(a_1, a_1)$ $\neg R$

~~Ancestor(Paul, John)~~

$\neg \text{Father}(a_1, a_1) \lor \text{Ancestor}(a_1, a_1)$ $\neg \text{Father} \lor \text{Ancestor}$

CDCL

$|N_{\text{BS}}| \simeq |N_{\text{SAT}}|$

- no instantiation overhead
- no redundant inferences
- short proofs
- compact models

BS to SAT one Variable

Finitely many constants: $a_1, a_2, \ldots, a_n$ n fixed

~~$\neg R(x, y) \vee \neg R(y, z) \vee R(x, z)$~~

$\neg R(x, x)$

$\text{Ancestor}(\text{Paul}, \text{John})$

~~$\neg \text{Father}(x, y) \vee \text{Ancestor}(x, y)$~~

$\neg R(a_1, a_1)$

$\vdots$

$\neg R(a_n, a_n)$

$\text{Ancestor}(\text{Paul}, \text{John})$

CDCL

$$n \cdot |N_{\text{BS}}| \simeq |N_{\text{SAT}}|$$

- instantiation overhead
- no redundant inferences
- long proofs
- large models

BS to SAT in General

Finitely many constants: $a_1, a_2, \ldots, a_n$ $\quad$ n fixed

$\neg R(x,y) \vee \neg R(y,z) \vee R(x,z)$	n^3 clauses
$\neg R(x,x)$	n clauses
$\mathrm{Ancestor}(\mathrm{Paul}, \mathrm{John})$	$\mathrm{Ancestor}(\mathrm{Paul}, \mathrm{John})$
$\neg\,\mathrm{Father}(x,y) \vee \mathrm{Ancestor}(x,y)$	n^2 clauses

CDCL

$$n^{|\mathsf{vars}|} \cdot |N_{\mathsf{BS}}| \;\simeq\; |N_{\mathsf{SAT}}|$$

- upfront instantiation overhead
- no redundant inferences
- long proofs
- large models

Ontology: $1M$ Clauses, $20M$ Facts, $10M$ Constants

BS via InstGen

$$N = \{\ldots, P(a) \lor Q(b), \ldots, \neg P(x_1) \lor R(x_1, x_2), \ldots\}$$

$$\Downarrow \quad \sigma = \{\vec{x} \mapsto b\}$$

iProver

$$N = \{\ldots, P(a) \lor Q(b), \ldots, \neg P(b) \lor R(b, b), \ldots\} \quad \Rightarrow \bot$$

$$\Downarrow$$

$$[\ldots, P(a), \ldots, \neg P(b), \ldots]$$

$$\Downarrow$$

$$N \cup \{\neg P(a) \lor R(a, x_2)\}$$

- no upfront instantiation overhead
- redundant inferences
- long proofs
- large models

[Ganzinger, Korovin]

BS via NRCL

$$N = \{\ldots, P(a) \lor Q(b), \ldots, \neg P(x_1) \lor R(x_1, x_2), \ldots\}$$

$$\Rightarrow^*_{\text{NRCL}} ([\ldots, P(a)^k], k, \top)$$

$$\Rightarrow_{\text{NRCL}} ([\ldots, P(a)^k, R(a, x_2)^{\neg P(x_1) \lor R(x_1,x_2)}], k, \top)$$

$$\Rightarrow_{\text{NRCL}} ([\ldots, P(a)^k, \neg P(x_1)^{k+1} : x_1 \neq a], k+1, \top)$$

Literal : $\bigwedge x_i \neq t_i$

t_i are constants, variables

emptiness NP-complete

- no instantiation
- no redundant inferences
- short proofs
- possibly compact models
- non-linear computations

[Alagi, Weidenbach]

BS via Superposition

Case Study: YAGO Ontology

YAGO Ontology: $1M$ Clauses, $20M$ Facts, $10M$ Constants

$\text{Ancestor}(\text{Paul}, \text{John})$

spass

$\neg R(x,y) \vee \neg R(y,z) \vee R(x,z)$

$x \succ y$

$\neg S(x) \vee \neg T(x) \vee Q(x)$

$S(x), T(x) \parallel Q(x)$

Satisfiability in 30min

Answer Queries in the range seconds

- no instantiation
- redundant inferences
- short proofs
- implicit model
- possibly huge clause sets
- Expert Set-Up

[SudaWischnewskiWeidenbach]

Bibliography

Superposition:

- [BachmairGanzinger 1994]
- [Handbook of Automated Reasoning 2001: [BachmairGanzinger], [NieuwenhuisRubio], [Weidenbach]]
- Baumgartner, Peltier, Suda, Voronkov, Waldmann, Weidenbach

Bernays-Schoenfinkel:

- [BaumgartnerFuchsTinelli 2006] [GanzingerKorovin 2003] [PiskacMouraBjorner 2010] [ClaessenSrensson 2003] [BonacinaPlaisted 2016] [PerezVoronkov 2008]

Conclusion

“FOL (Model-Based) Reasoning is Difficult,”

“and sometimes it is not the Right Approach,”

“and there is no single Right Approach.”

Applications: Expert Automation Support (Isabelle)

Thanks for Your Attention