Combinations of Theories for Decidable Fragments of First-order Logic

Pascal Fontaine

Loria, INRIA, Université de Nancy (France)

FroCoS, Trento
September 17, 2009

Context / Motivation

The veriT solver

www.verit-solver.org

- Satisfiability Modulo Theories SMT
- Combination of theories: uninterpreted symbols, arithmetic
- Satisfiability checking for formulas like
  $a \leq b \wedge b \leq a + x \wedge x = 0 \wedge \left[f(a) \neq f(b) \vee (p(a) \wedge \neg p(b+x)) \right]$
- Proof obligations for verification of distributed algorithm: B, TLA+ specifications
- Extend the language with operators for sets, relations,. . .

Introducing sets: operators

SMT + Syntactic sugar:

operator	Definition
$\in$	$\lambda xp.\ p(x)$
$\cap$	$\lambda pq.\ \lambda x.\ p(x) \wedge q(x)$
$\setminus$	$\lambda pq.\ \lambda x.\ p(x) \wedge \neg q(x)$
$\subseteq$	$\lambda pq.\ \forall x.\ p(x) \Rightarrow q(x)$
$\vdots$	$\vdots$
transitive	$\lambda r.\ \forall xyz.\ [r(x,y) \wedge r(y,z)] \Rightarrow r(x,z)$
$\vdots$	$\vdots$
permutation	$\lambda r.\ \forall xyz.\ r(x,y,z) = r(y,z,x) = r(z,x,y)$

- introduces quantifiers
- sat. checking in combination of initial theories + FOL theory

Introducing sets: an example

For example :

$$a = b \wedge (\{f(a)\} \cup E) \subseteq A \wedge f(b) \notin C \wedge A \cup B = C \cap D$$

becomes

$$\begin{aligned} a = b \wedge \forall x[(x = f(a) \vee E(x)) \Rightarrow A(x)] \wedge \neg C(f(b)) \\ \wedge \forall x.\ [A(x) \vee B(x)] \equiv [C(x) \wedge D(x)] \end{aligned}$$

- quantifiers come from second-order equalities, operators that contain quantifiers
- but the obtained FOL theory is BSR: $\exists^*\forall^*\varphi$ (φ function- and quantifier-free), and (for sets) monadic

Motivation - problem - solution

- Motivation: extend the language of SMT solvers with operators on sets, relations,. . .
- Problem: combine a Bernays-Schönfinkel-Ramsey theory with a decidable fragment (the initial language of the SMT solver)

It is indeed possible to combine a decidable theory from the BSR, monadic, or two variable classes, with (nearly) any decidable theory

FOL decidable classes and combinations

SMT solvers:

- satisfiability checking of (quantifier-free) formulas in a static combination of theories
- theories: disjoint, FOL, equational, decidable, stably infinite
- e.g. empty theory, linear arithmetic, arrays, lists, bitvectors

Some major decidable equational FOL theories:

- Bernays-Schönfinkel-Ramsey: $\exists^*\forall^*\varphi$ (φ function- and quantifier-free)
- two-variables relational fragment
- monadic first-order logic

Those theories are not stably infinite: $\forall x \forall y\, x = y$
Nelson-Oppen not applicable

Combining disjoint decision procedures (1)

A combination of disjoint languages:

$$L = \{x \leq y,\, y \leq x + f(x),\, P(h(x) - h(y)),\, \neg P(0),\, f(x) = 0\}$$

uninterpreted symbols (P, f, h), and arithmetic ($+$, $-$, $\leq$, 0).

Combination of disjoint decision procedures

Combination of the empty theory and theory for linear arithmetic (both stably-infinite)

Separation using new variables:

$$\begin{aligned} L_1 &= \{x \leq y,\, y \leq x + v_1,\, v_1 = 0,\, v_2 = v_3 - v_4,\, v_5 = 0\} \\ L_2 &= \{P(v_2),\, \neg P(v_5),\, v_1 = f(x),\, v_3 = h(x),\, v_4 = h(y)\}. \end{aligned}$$

L and $L_1 \cup L_2$ both satisfiable or both unsatisfiable.

Combining disjoint decision procedures (2)

Cooperation by exchanging equalities:

$$\begin{aligned} L_1 &= \{x \leq y,\, y \leq x + v_1,\, v_1 = 0, v_2 = v_3 - v_4,\, v_5 = 0\} \\ L_2 &= \{P(v_2),\, \neg P(v_5),\, v_1 = f(x),\, v_3 = h(x),\, v_4 = h(y)\} \end{aligned}$$

From L_1, $x = y$:

$$\begin{aligned} L_1 &= \{x \leq y,\, y \leq x + v_1,\, v_1 = 0,\, v_2 = v_3 - v_4,\, v_5 = 0\} \\ L_2' &= \{P(v_2),\, \neg P(v_5),\, v_1 = f(x),\, v_3 = h(x),\, v_4 = h(y),\, x = y\} \end{aligned}$$

From L_2', $v_3 = v_4$:

$$\begin{aligned} L_1' &= \{x \leq y,\, y \leq x + v_1,\, v_1 = 0,\, v_2 = v_3 - v_4,\, v_5 = 0, v_3 = v_4\} \\ L_2' &= \{P(v_2),\, \neg P(v_5),\, v_1 = f(x),\, v_3 = h(x),\, v_4 = h(y),\, x = y\} \end{aligned}$$

From L_1', $v_2 = v_5$:

$$\begin{aligned} L_1' &= \{x \leq y,\, y \leq x + v_1,\, v_1 = 0,\, v_2 = v_3 - v_4,\, v_5 = 0, v_3 = v_4\} \\ L_2'' &= \{P(v_2),\, \neg P(v_5),\, v_1 = f(x),\, v_3 = h(x),\, v_4 = h(y),\, x = y, v_2 = v_5\} \end{aligned}$$

L_2'' is unsatisfiable.

Combining disjoint decision procedures (2)

Cooperation by exchanging equalities:

$$L_1 = \{x \leq y,\ y \leq x + v_1,\ v_1 = 0, v_2 = v_3 - v_4,\ v_5 = 0\}$$
$$L_2 = \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y)\}$$

From L_1, $x = y$:

$$L_1 = \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0\}$$
$$L_2' = \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y\}$$

From L_2', $v_3 = v_4$:

$$L_1' = \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0, v_3 = v_4\}$$
$$L_2' = \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y\}$$

From L_1', $v_2 = v_5$:

$$L_1' = \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0, v_3 = v_4\}$$
$$L_2'' = \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y, v_2 = v_5\}$$

L_2'' is unsatisfiable.

Combining disjoint decision procedures (2)

Cooperation by exchanging equalities:

$$\begin{aligned} L_1 &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0, v_2 = v_3 - v_4,\ v_5 = 0\} \\ L_2 &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y)\} \end{aligned}$$

From L_1, $x = y$:

$$\begin{aligned} L_1 &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0\} \\ L_2' &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y\} \end{aligned}$$

From L_2', $v_3 = v_4$:

$$\begin{aligned} L_1' &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0, v_3 = v_4\} \\ L_2' &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y\} \end{aligned}$$

From L_1', $v_2 = v_5$:

$$\begin{aligned} L_1' &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0, v_3 = v_4\} \\ L_2'' &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y, v_2 = v_5\} \end{aligned}$$

L_2'' is unsatisfiable.

Combining disjoint decision procedures (2)

Cooperation by exchanging equalities:

$$\begin{aligned} L_1 &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0, v_2 = v_3 - v_4,\ v_5 = 0\} \\ L_2 &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y)\} \end{aligned}$$

From L_1, $x = y$:

$$\begin{aligned} L_1 &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0\} \\ L_2' &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y\} \end{aligned}$$

From L_2', $v_3 = v_4$:

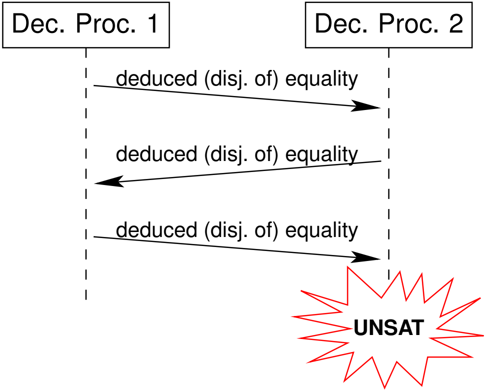
$$\begin{aligned} L_1' &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0, v_3 = v_4\} \\ L_2' &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y\} \end{aligned}$$

From L_1', $v_2 = v_5$:

$$\begin{aligned} L_1' &= \{x \leq y,\ y \leq x + v_1,\ v_1 = 0,\ v_2 = v_3 - v_4,\ v_5 = 0, v_3 = v_4\} \\ L_2'' &= \{P(v_2),\ \neg P(v_5),\ v_1 = f(x),\ v_3 = h(x),\ v_4 = h(y),\ x = y, v_2 = v_5\} \end{aligned}$$

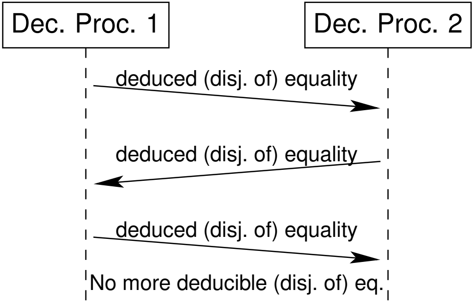
L_2'' is unsatisfiable.

Combining disj. DPs : “unsatisfiable” scenario



Sound : every deduced fact is a consequence of the original set of formulas

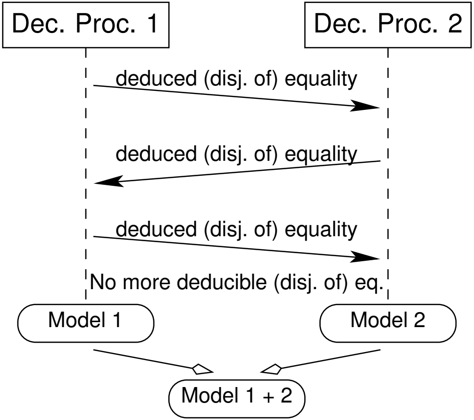
Combining disj. DPs : “satisfiable” scenario



Really SAT? (Complete?)

- all disjunctions of equalities propagated
- models agree on cardinalities

Combining disj. DPs : "satisfiable" scenario



Really SAT? (Complete?)

- all disjunctions of equalities propagated
- **models agree on cardinalities**

Ensuring agreement on cardinalities?

Different frameworks (and capabilities)

- Nelson-Oppen:
  requirement on theories: stably infinite (not suitable for BSR)
  if satisfiable, there is an infinite model (FOL theories $\Rightarrow \aleph_0$)
- Combining with the empty theory (and some others):
  the empty theory does not constraint much the cardinalities
- ...

Cardinalities and decidable fragments

Decidable classes

- Bernays-Schönfinkel-Ramsey: $\exists^*\forall^*\varphi$ (φ function- and quantifier-free)
- two-variables relational fragment
- monadic first-order logic

all have following property (pumping theorem)

for every theory $\mathcal{T}$, there is a computable $k(\mathcal{T})$ s. t. if there is a model of cardinality $\geq k(\mathcal{T})$, there is a model of every cardinality $\geq k(\mathcal{T})$.

The set of cardinalities is the finite or cofinite set:

$$S_{\mathcal{T}} \cup \left\{ \kappa \mid \kappa \text{ is a cardinality} \wedge \kappa \geq k(\mathcal{T}) \right\}$$

with $S_{\mathcal{T}} \subset \mathbb{N}$ computable and finite, and $k(\mathcal{T})$ computable ($\mathcal{T}$ is gentle).

Cardinalities and decidable fragments (2)

Pumping theorem:

for every theory $\mathcal{T}$, there is a computable $k(\mathcal{T})$ s. t. if there is a model of cardinality $\geq k(\mathcal{T})$, there is a model of every cardinality $\geq k(\mathcal{T})$.

For instance, $\mathcal{T}$ is a Löwenheim theory (other classes are “similar”)

- assume there is no constant in $\mathcal{T}$ (can be relaxed)
- n is the number of predicates
- q is the number of imbricated quantifiers
- there is 2^n different configurations (tables, types) for elements of the domain with respect to the n predicates
- if there exists a model with cardinality $\geq q\,2^n$ then there should be $\geq q$ elements with the same configuration
- any such element can be duplicated, to infinity
- proved by induction on the structure of formulas in $\mathcal{T}$

Combination “in practice”

While combining a BSR, Monadic, or 2-variables theory $\mathcal{T}_1$ with another theory $\mathcal{T}_2$

- first propagate all (disjunctions of) equalities
- if still satisfiable, compute the set of cardinalities for $\mathcal{T}_1 \cup L_1$
- if the set is finite, check every cardinality against $\mathcal{T}_2 \cup L_2$
- if the set is infinite,
  - check every cardinality $< k$ against $\mathcal{T}_2 \cup L_2$
  - check if $\mathcal{T}_2 \cup L_2$ accepts a cardinality $\geq k$ by checking the satisfiability of $\mathcal{T}_2 \cup L_2 \cup \{a_i \neq a_j \mid 0 < i, j \leq k\}$ where a_is are new constants
- if one cardinality is acceptable for $\mathcal{T}_2 \cup L_2$, then the original problem is satisfiable. Otherwise it is not.

Conclusion and future works

- veriT includes FOL ATP (currently E, also Spass in the future)
- Saturation provers are (or can be turned into) decision procedures for decidable FOL fragments
- Long term goal: raise the degree of completeness of the combination SMT+FOL

Future works:

- is there any other interesting suitable decidable fragment? The guarded fragment?
- how can we really turn this into something usable? Negotiation of cardinality