Variant-Based Decidable Satisfiability in Initial Algebras with Predicates

Raúl Gutiérrez\textsuperscript{1}  José Meseguer\textsuperscript{2}

\textsuperscript{1}DSIC, Universitat Politècnica de València, Spain

\textsuperscript{2}University of Illinois at Urbana-Champaign, Illinois, USA

\textbf{Namur (Belgium), October 11, 2017}
Motivation

1. Some of the most recent advances in software verification are due to the systematic use of decision procedures in model checkers and theorem provers.

2. For a system specified by theory $T$, SMT solving can partially automate verification by using procedures for decidable subtheories $T_i$.

3. Limitation of SMT tools: lack of extensibility of decidable fragment.

4. Users can extend a specification’s decidable fragment if theory-generic decision procedures are added.

5. Variant-based satisfiability (VS): a decision procedure for initial algebras $T_{\Sigma/E\cup B}$ generic on theories $(\Sigma, E \cup B)$ under quite general conditions.

6. Limitation: current VS algorithm applies well to user-definable data structures, but cannot handle user-definable predicates.
Extend variant-based satisfiability to initial algebras with user-definable predicates under fairly general conditions using two key ideas:

1. characterizing the cases when $p(u_1, \ldots, u_n) \neq tt$ by means of constrained patterns; and

2. eliminating all occurrences of disequalities of the form $p(u_1, \ldots, u_n) \neq tt$ in a quantifier-free (QF) formula by means of such patterns.
Outline

1 Motivation
2 Variant Satisfiability
3 Predicates
4 OS-compactness
5 Negative Patterns
6 Inductive Satisfiability Decision Procedure
7 Implementation
8 Conclusions
Example: Sets of Natural Numbers \((\Sigma, E \cup B)\)

\[
\text{fmod ACU-NAT is}
\]
\[
\text{sort Natural .}
\]
\[
\text{op 0 : } \rightarrow \text{Natural [ctor] .}
\]
\[
\text{op 1 : } \rightarrow \text{Natural [ctor] .}
\]
\[
\text{op \_\_+_ : Natural Natural } \rightarrow \text{Natural [ctor assoc comm id: 0] .}
\]
\[\text{endfm}\]

\[
\text{fmod ACU-NAT-FUN is}
\]
\[
\text{pr ACU-NAT .}
\]
\[
\text{op max : Natural Natural } \rightarrow \text{Natural [comm] .}
\]
\[
\text{op min : Natural Natural } \rightarrow \text{Natural [comm] .}
\]
\[
\text{op \_-_- : Natural Natural } \rightarrow \text{Natural .}
\]
\[\text{*** monus}
\]
\[
\text{vars N M : Natural .}
\]
\[
\text{eq max(N,N + M) = N + M [variant] .}
\]
\[
\text{eq min(N,N + M) = N [variant] .}
\]
\[
\text{eq N - (N + M) = 0 [variant].}
\]
\[
\text{eq (N + M) - N = M [variant] .}
\]
\[\text{endfm}\]
Example: Sets of Natural Numbers \((\Sigma, E \cup B)\)

```latex
fmod ACU-NAT-SET is
  pr ACU-NAT.

  sort NaturalSet.
  sort Pred.

  subsort Natural < NaturalSet.

  op mt : -> NaturalSet [ctor].
  op _,_ : NaturalSet NaturalSet -> NaturalSet [ctor assoc comm].
  op tt : -> Pred [ctor].
  *** set containment
  op _=C_ : NaturalSet NaturalSet -> Pred [ctor].

  vars NS NS' : NaturalSet.

  *** identity of set union
  eq NS , mt = NS [variant].
  *** idempotency of set union
  eq NS , NS = NS [variant].
  *** idempotency of set union
  eq NS , NS , NS' = NS , NS' [variant].
  eq mt =C NS = tt [variant].
  eq NS =C NS = tt [variant].
  eq NS =C NS , NS' = tt [variant].

endfm
```

R. Gutiérrez & J. Meseguer (UPV & UIUC)
Variants

Given a decomposition $\mathcal{R} = (\Sigma, B, \bar{E})$ of a MS equational theory $(\Sigma, E)$ and a $\Sigma$-term $t$, a variant of $t$ is a pair $(u, \theta)$ such that:

- $u =_B (t\theta)!_{\bar{E},B}$,
- $\text{dom}(\theta) \subseteq \text{vars}(t)$, and
- $\theta = \theta!_{\bar{E},B}$, that is, $\theta(x) = \theta(x)!_{\bar{E},B}$ for all variables $x$. $(u, \theta)$ is called a ground variant iff, furthermore, $u \in T_\Sigma$.

Given variants $(u, \theta)$ and $(v, \gamma)$ of $t$, $(u, \theta)$ is called more general than $(v, \gamma)$, denoted $(u, \theta) \sqsupseteq_B (v, \gamma)$, iff there is a substitution $\rho$ such that:

- $(\theta \rho)|_{\text{vars}(t)} =_B \gamma$, and
- $u \rho =_B v$.

Let $[t]_{\bar{E},B} = \{(u_i, \theta_i) \mid i \in I\}$ denote a complete set of variants of $t$, that is, a set of variants such that for any variant $(v, \gamma)$ of $t$ there is an $i \in I$, such that $(u_i, \theta_i) \sqsupseteq_B (v, \gamma)$. 
Example: Variants

get variants in ACU-NAT-FUN:
min(1, N:Natural + K:Natural) .

Variant #1
Natural: min(1, N:Natural + K:Natural)

Variant #2
Natural: 1
K:Natural --> 1 + K1:Natural

Variant #3
Natural: 1
N:Natural --> 1 + N1:Natural

Variant #4
Natural: 0
N:Natural --> 0
K:Natural --> 0

get variants in ACU-NAT-FUN:
N:Natural - K:Natural .

Variant #1
Natural: N:Natural - K:Natural

Variant #2
Natural: 0
K:Natural --> K1:Natural + N:Natural

Variant #3
Natural: N1:Natural
N:Natural --> N1:Natural + K:Natural

Variant #4
Natural: 0
N:Natural --> 0
K:Natural --> 0
Finite Variant Property

- A decomposition $\mathcal{R} = (\Sigma, B, R)$ has the finite variant property (FVP) iff for each $\Sigma$-term $t$ there is a finite complete set of variants $\llbracket t \rrbracket_{R,B} = \{(u_1, \theta_1) \ldots (u_n, \theta_n)\}$.

- If $B$ has a finitary $B$-unification algorithm, and $\mathcal{R} = (\Sigma, B, R)$ has FVP, $\llbracket t \rrbracket_{R,B}$ can be chosen to be the set of most general variants.

Note

FVP easy to check when it holds. Example: ACU-NAT-SET is FVP.
Representing Predicates

- A **predicate** is viewed as a function symbol $p : s_1 \ldots s_n \to \text{Pred}$, with \text{Pred} a new sort having constant \text{tt}.
- An atomic formula $p(t_1, \ldots, t_n)$ is then expressed as the equation $p(t_1, \ldots, t_n) = \text{tt}$. 
Example: Predicates on Sets of Natural Numbers

fmod ACU-NAT-SET-PRED is
  pr ACU-NAT-SET .

  *** strict order
  op _>_ : Natural Natural -> Pred [ctor] .

  *** sort predicates
  op natural : NaturalSet -> Pred [ctor] .
  op even : NaturalSet -> Pred [ctor] .
  op odd : NaturalSet -> Pred [ctor] .

  vars N M : Natural .

  eq N + M + 1 > N = tt [variant] .

  eq natural(N) = tt [variant] .

  eq even(N + N) = tt [variant] .

  eq odd(N + N + 1) = tt [variant] .
endfm
**Constructor Variants**

**Question**
What variants of t cover as instances modulo B all canonical forms of all ground instances of t?

Let $R = (\Sigma, B, R)$ be an FVP decomposition of $(\Sigma, E)$ protecting a constructor decomposition $R_\Omega = (\Omega, B_\Omega, R_\Omega)$. Assume that:

- $\Sigma = \Omega \cup \Delta$ with $\Omega \cap \Delta = \emptyset$;
- $B$ has a finitary $B$-unification algorithm and $B = B_\Omega \cup B_\Delta$, with $B_\Omega$ $\Omega$-equations and if $u = v \in B_\Delta$, $u,v$ are non-variable $\Delta$-terms.

Call $\llbracket t \rrbracket_{R,B}^\Omega = \{(v, \theta) \in \llbracket t \rrbracket_{R,B} : v \in T_\Omega(X)\}$ the set of constructor variants of t.

**Answer**
If $[u] \in C_{R_\Omega}$ is of the form $u \equiv_B (t\gamma)!_{R,B}$, then there is $(v, \theta) \in \llbracket t \rrbracket_{R,B}^\Omega$ and a normalized ground substitution $\tau$ such that $u \equiv_B v\tau$. 

R. Gutiérrez & J. Meseguer (UPV & UIUC)
An equational OS-FO theory \((\Sigma, E)\) is called **OS-compact** iff:

- for each sort \(s\) in \(\Sigma\) we can effectively determine whether \(s\) is finite or infinite in \(T_{\Sigma/E,s}\), and, if finite, can effectively compute a representative ground term \(\text{rep}([u]) \in [u]\) for each \([u] \in T_{\Sigma/E,s}\);
- \(=_{E}\) is decidable and \(E\) has a finitary unification algorithm; and
- any finite conjunction \(\bigwedge D\) of negated \(\Sigma\)-atoms whose variables all have infinite sorts and such that \(\bigwedge D\) is \(E\)-consistent is satisfiable in \(T_{\Sigma,E}\).

Call an OS theory \((\Sigma, E)\) **OS-compact** iff OS-FO theory \((\Sigma, E)\) is OS-compact.

**Theorem**

If \((\Sigma, E)\) is an **OS-compact** theory, then satisfiability of QF \(\Sigma\)-formulas in \(T_{\Sigma,E}\) is decidable.
Current Variant Satisfiability

Theorem 1
If \((\Omega, B_\Omega)\) has \(B_\Omega\) only with \(ACCU\)-axioms, then \((\Omega, B_\Omega)\) is OS-compact.

Theorem 2 (Variant Satisfiability)
If \((\Sigma, E \cup B)\) if FVP and protects \((\Omega, B_\Omega)\) with \(B_\Omega \subseteq ACCU\), then QF satisfiability in \((\Sigma, E \cup B)\) is decidable.
Limitation

Question

What happens with the user-definable predicates?

- $p$ is a constructor operator of sort $\text{Pred}$ which is not a free constructor modulo the axioms $B_\Omega$.
- The OS-compactness of a constructor decomposition $\mathcal{R}_\Omega = (\Omega, B_\Omega, R_\Omega)$ can be broken (or be a hard to prove task) when adding user-definable predicates.

Solution

We provide a decision procedure for validity and satisfiability of QF formulas in the initial algebra of an FVP theory $\mathcal{R}$ that may contain user-definable predicates and protects a constructor decomposition $\mathcal{R}$ that need not be OS-compact under reasonable assumptions.
Example: Negative Patterns

- Greater than: $N > N + M$
- Even:
  - $\text{even}(mt)$
  - $\text{even}(N + N + 1)$
  - $((N = C \ NS /= tt), (NS /= mt)) \implies \text{even}((N, NS))$
- Odd:
  - $\text{odd}(mt)$
  - $\text{odd}(N + N)$
  - $((N = C \ NS /= tt), (NS /= mt)) \implies \text{odd}((N, NS))$
- Natural:
  - $\text{natural}(mt)$
  - $((N = C \ NS /= tt), (NS /= mt)) \implies \text{natural}((N, NS))$
Negative Patterns

- Negative constrained patterns are of the form:

\[
\bigwedge_{1 \leq l \leq n_j} w^j_l \neq w'^j_l \Rightarrow p(v^j_1, \ldots, v^j_n) \neq tt, \quad 1 \leq j \leq m_p
\]

with the \(v^j_i\), \(w^j_l\) and \(w'^j_l\) \(\Omega_c\)-terms with variables in \(Y_j = \text{vars}(p(v^j_1, \ldots, v^j_n))\).

- These negative constrained patterns are interpreted as meaning that the following semantic equivalences are valid in \(C_R\) for each \(p \in \Omega_\Pi\), where \(\rho_j \in \{\rho \in [Y_j \rightarrow T_{\Omega_c}] \mid \rho = \rho!_{R,B}\}\), \(B = B_\Delta \uplus B_{\Omega_c}\), and \(R = R_\Delta \uplus R_{\Omega_c} \uplus R_\Pi\):

\[
[p(v^j_1, \ldots, v^j_n)\rho_j] \in C_R \Leftrightarrow \bigwedge_{1 \leq l \leq n_j} (w^j_l \neq w'^j_l)\rho_j
\]

\[
[p(t_1, \ldots, t_n)] \in C_R \Leftrightarrow \exists j \exists \rho_j \ [p(t_1, \ldots, t_n)] = [p(v^j_1, \ldots, v^j_n)\rho_j] \land \bigwedge_{1 \leq l \leq n_j} (w^j_l \neq w'^j_l)\rho_j
\]
The inductive validity decision problem of whether $C_R \models \varphi$ is reduced to deciding whether $\neg \varphi$ is unsatisfiable in $C_R$.

In this way, it is enough to decide the satisfiability of a conjunction of $\Sigma$-litersals of the form $\bigwedge G \land \bigwedge D$ (the QF $\Sigma$-formula in disjunctive normal form), where the $G$ are equations and the $D$ are disequations.

Steps:

1. **Unification.** Satisfiability of the conjunction $\bigwedge G \land \bigwedge D$ is replaced by satisfiability for some conjunction in the set $\{ (\bigwedge D\alpha)_{R,B} \mid \alpha \in \text{VarUnif}_E(\bigwedge G) \}$. 
The Inductive Satisfiability Decision Procedure

(2/2)

2 **Π-Elimination.** For each $\bigwedge D' = \bigwedge D_1 \land p(t_1, \ldots, t_n) \neq tt \land \bigwedge D_2$, we replace $\bigwedge D'$ by all not obviously unsatisfiable conjunctions of the form:

$$\left( \bigwedge D_1 \land \bigwedge_{1 \leq l \leq n_j} w^{j_l} \neq \alpha^{j_l} \land \bigwedge D_2 \right) \theta \alpha$$

where $1 \leq j \leq m_p$, $W = \text{vars}(\bigwedge D')$, $(p(t'_1, \ldots, t'_n), \theta) \in \llbracket p(t_1, \ldots, t_n) \rrbracket_{R, B}^W$, and $\alpha$ is a disjoint $B_{\Omega_c}$-unifier of the equation $p(t'_1, \ldots, t'_n) = p(v^{j_1}_1, \ldots, v^{j_n}_n)$.

3 **Reduce Conjunctions of \(\Sigma\) Disequalities to Conjunctions of \(\Omega_c\) Disequalities.** For $\bigwedge D'$ a $\Delta \cup \Omega_c$-conjunction of disequalities, viewed as a ($\Delta \cup \Omega_c$)-term its constructor $\Omega_c^\wedge$-variants are of the form $(\bigwedge D'', \gamma)$, with $\bigwedge D''$ an $\Omega_c$-conjunction of disequalities. Then $\bigwedge D'$ is satisfiable in $C_R$ iff some $\bigwedge D'' \tau$ so obtained is $B_{\Omega_c}$-consistent for some $\Omega_c^\wedge$-variant $(\bigwedge D'', \gamma)$ of $\bigwedge D'$. 
We have implemented the variant satisfiability decision procedure in a new prototype tool.

The implementation consists of 11 new Maude modules (from 17 in total), 2345 new lines of code, and uses the Maude’s META-LEVEL to carry out the steps of the procedure in a reflective way.

We have also developed a Maude interface to ease the definition of properties and patterns as equations. The three steps of the variant satisfiability procedure are implemented using Maude’s META-LEVEL functions.
Example: Odd and Even

mod ACU-NAT-SET-PRED-CONJECTURES is
    pr ACU-NAT-SET-PRED-PATTERNS .

*** odd(N) = tt \iff even(N) /= tt .
    op prop1 : Natural -> AtomMagma .
    op prop2 : Natural -> AtomMagma .

eq prop1(N)
    = (odd(N) = tt) , (even(N) = tt) .
eq prop2(N)
    = (even(N) /= tt) , (odd(N) /= tt) .
endm

Unification of prop1:
    No variant unifiers can be found.

Unification of prop2:
    (even(N) /= tt) , (odd(N) /= tt)

Predicate elimination of prop2:
    even(M + M) /= tt , odd(M + M) /= tt \Rightarrow
    tt /= tt , odd(M + M) /= tt

Unsatisfiable!
Example: Greater Than

mod ACU-NAT-SET-PRED-CONJECTURES is
pr ACU-NAT-SET-PRED-PATTERNS .

*** N > M = tt \/ N = M \/ M > N = tt

op prop : Natural Natural -> AtomMagma .

eq prop(N,M)
  = (N > M /= tt) ,
    (N /= M) ,
    (M > N /= tt) .
endm

Unification of prop:
(N > M /= tt) ,
(N /= M) ,
(M > N /= tt)

Predicate elimination of prop:
(N > N + 0 /= tt) ,
(N /= N + 0) ,
(N + 0 > N /= tt) =>
(N /= N)

Unsatisfiable!
Conclusions and future work

- Satisfiability decision procedures can be either theory-specific or theory-generic. These two classes of procedures complement each other: theory specific ones are more efficient; but theory-generic ones are user-definable and can substantially increase the range of SMT solvers.

- Our work has extended variant satisfiability to support initial algebras specified by FVP theories with user-definable predicates under fairly general conditions. Since such predicates are often needed in specifications, this substantially enlarges the scope of variant-based initial satisfiability algorithms.

- The most obvious next step is to combine the original variant satisfiability algorithm with the present one.

- Furthermore, our goal is to include this powerful decision procedure in our automatic inductive theorem prover $\nu$-ITP.