Decision Procedures and Verification

Martin Blicha

Charles University

14.5.2018

Combination of theories

Introduction

- Decision procedures seen so far focus on specific theory.
- Often formulas generated from verification conditions mix expressions from several theories.
  - Most prominent example is linear arithmetic and uninterpreted functions.
- Combination of decision procedures for involved theories to obtain decision procedure for the combination.
  - Nelson–Oppen combination method
    - Nelson, Oppen, *Simplification by cooperating decision procedures, 1979*
  - Delayed Theory Combination
    - Bozzano at al., *Efficient Satisfiability Modulo Theories via Delayed Theory Combination, 2005*
  - Model-based Theory Combination
    - de Moura, Bjørner, *Model-based Theory Combination, 2007*

Combination formally

- A theory is defined over a signature Σ
  - Set of non-logical symbols (predicate and function symbols).
- Theory T is a set of sentences.
  - More commonly represented by a set of *axioms*.
  - Theory is the set of sentences deriveable from the axioms.

Definition (theory combination)

Given two theories T_1 and T_2 with signatures Σ_1 and Σ_2, respectively, the theory combination $T_1 \oplus T_2$ is a $\Sigma_1 \cup \Sigma_2$-theory defined by the axiom set $T_1 \cup T_2$.

- *Theory combination problem* is to decide whether φ, a $\Sigma_1 \cup \Sigma_2$ formula, is $T_1 \oplus T_2$ valid.

Convex theory

Definition (convex theory)

A Σ-theory T is *convex* if for every conjunctive Σ-formula φ

$$(\varphi \implies \bigvee_{i=1}^{n} x_i = y_i) \text{ is } T\text{-valid for some finite } n > 1 \implies$$

$$(\varphi \implies x_i = y_i) \text{ is } T\text{-valid for some } i \in \{1, \ldots, n\},$$

where x_i, y_i are some variables.

- Linear arithmetic over reals is convex.
  - A conjunction of linear arithmetic predicates define either empty set, singleton or infinite set of values.
- Linear arithmetic over integers is not convex.
  - $x_1 = 1 \wedge x_2 = 2 \wedge 1 \leq x_3 \wedge x_3 \leq 2 \implies (x_3 = x_1 \vee x_3 = x_2)$

Nelson–Oppen restrictions

- Nelson-Oppen combination procedure solves combination problem for theories (under certain restrictions).

Definition (Nelson–Oppen restrictions)

In order for the Nelson–Oppen procedure to be applicable, the theories T_1, T_2 should comply with the following restrictions:

1. T_1, T_2 are quantifier-free first-order theories with equality.
2. There is a decision procedure for each of the theories.
3. The signatures of the theories are disjoint.
4. The theories are interpreted over infinite domain.

Purification

- Satisfiability-preserving transformation after which each atom is from a specific theory.
  - Afterwards, all atoms are *pure*.
- For a give formula φ, purification generates an equisatisfiable φ' the following way.
  1. Let $\varphi' := \varphi$.
  2. For each "alien" subexpression in t in φ'.
     2.1 Replace t with a new auxiliary variable a_t.
     2.2 $\varphi' := \varphi' \wedge a_t = t$.
- Example: $\varphi := x_1 \leq f(x_1) \implies \varphi' := x_1 \leq a \wedge a = f(x_1)$
- After purification, φ' can be partitioned to conjunctions of T_i-literals.

Nelson–Oppen procedures for convex theories

Algorithm NELSON–OPPEN-CONVEX

1. Purification: Purify φ into $F_1, \ldots, F_k$.
2. Apply the decision procedure for T_i to F_i. If there exists i such that F_i is unsatisfiable in T_i return UNSAT.
3. Equality propagation: If there exist i, j such that F_i T_i-implies an equality between variables of φ that is not T_j-implied by F_j, add this equality to F_j and go to step 2.
4. Return SAT.

- Example $(f(x_1, 0) \geq x_3) \wedge (f(x_2, 0) \leq x_3) \wedge (x_1 \geq x_2) \wedge (x_2 \geq x_1) \wedge (x_3 - f(x_1, 0) \geq 1$

Combining Nonconvex Theories

- Example where NELSON–OPPEN-CONVEX fails:
  - For linear arithmetic over integers and uninterpreted predicates.
  - $1 \leq x \wedge x \leq 2 \wedge p(x) \wedge \neg p(1) \wedge \neg p(2)$.

Combining Nonconvex Theories

- Example where NELSON–OPPEN-CONVEX fails:
  - For linear arithmetic over integers and uninterpreted predicates.
  - $1 \leq x \wedge x \leq 2 \wedge p(x) \wedge \neg p(1) \wedge \neg p(2)$.
- Remedy is to consider not only implied equalities, but also *disjunctions* of equalities.
  - There are finitely many of them (which are non-equivalent).

Combining Nonconvex Theories

- Example where NELSON–OPPEN-CONVEX fails:
  - For linear arithmetic over integers and uninterpreted predicates.
  - $1 \leq x \land x \leq 2 \land p(x) \land \neg p(1) \land \neg p(2)$.
- Remedy is to consider not only implied equalities, but also *disjunctions* of equalities.
  - There are finitely many of them (which are non-equivalent).
- Problem is split to as many parts as there are disjuncts and the procedure is called recursively.
  - In the example, the disjunction $x = 1 \lor x = 2$ is implied.

Nelson–Oppen Procedure For Nonconvex Theories

Algorithm $\textsc{Nelson–Oppen}$

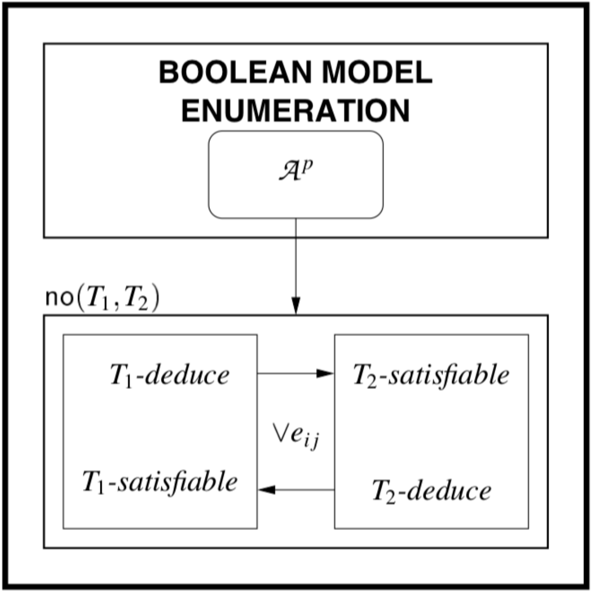
1. Purification: Purify φ into $F_1, \ldots, F_k$.
2. Apply the decision procedure for T_i to F_i. If there exists i such that F_i is unsatisfiable in T_i return UNSAT.
3. Equality propagation: If there exist i, j such that F_i T_i-implies an equality between variables of φ that is not T_j-implied by F_j, add this equality to F_j and go to step 2.
4. Splitting: If there exists i such that
   - $F_i \implies (x_1 = y_1 \vee \cdots \vee x_k = y_k)$ and
   - $\forall j \in \{1, \ldots, k\}. F_i \;\not\Longrightarrow\; x_j = y_j$,

   then apply $\textsc{Nelson–Oppen}$ recursively to $\varphi' \wedge x_1 = y_1, \ldots, \varphi' \wedge x_k = y_k$. If any of these subproblems is satisfiable, return SAT, otherwise return UNSAT.
5. Return SAT.

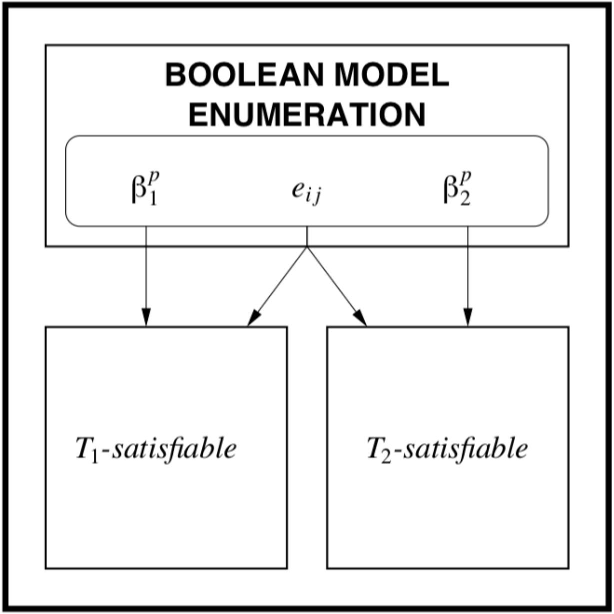
Deficiencies of Nelson–Oppen Procedure

- Let dp_1, dp_2 denote the decision procedures for T_1, T_2 and $no(T_1, T_2)$ denote the Nelson–Oppen procedure for $T_1 \oplus T_2$.
- In DPLL(T) framework *no* works as a single decision procedure.
- Additional requirements imposed on individual decision procedures dp_1 and dp_2:
  - Deduction of (disjunctions of) equalities.
  - Mutual awareness and comunication interface (for exchanging equalities).

Nelson–Oppen procedure in DPLL(T) framework



Delayed Theory Combination



Delayed Theory Combination

- Does *not* require direct combination of T_1 and T_2.
- dp_1, dp_2 communicate only with SAT solver (Boolean enumerator of assignments)
  - No deduction of equalities is needed.
- Consistency is assured by introduction of *interface equalities* to the Boolean skeleton of input formula.
  - Interface variable = variable that is common to both parts of the purified formula.
- Both theory solvers get the same assignment for interface equalities.
- This ensures that the partial models can be merged to single model for the input formula.

Delayed Theory Combination - algorithm

procedure $\textsc{Delayed-Theory-Combination}(\varphi)$
    $\psi \leftarrow \textsc{Purify}(\varphi)$
    $\mathcal{A}^p \leftarrow \mathit{fol2prop}(\mathit{Atoms}(\psi) \cup E(\mathit{interface\_vars}(\psi)))$
    $\psi^p \leftarrow \mathit{fol2prop}(\psi)$
    while Bool-satisfiable(ψ^p) **do**
        $\beta_1^p \wedge \beta_2^p \wedge \beta_e^p = \beta^p \leftarrow \mathit{total\_assingment}(\mathcal{A}^p, \psi^p)$
        $(\rho_1, \pi_1) \leftarrow T_1$-satisfiable (prop2fol$(\beta_1^p \wedge \beta_e^p)$)
        $(\rho_2, \pi_2) \leftarrow T_2$-satisfiable (prop2fol$(\beta_2^p \wedge \beta_e^p)$)
        if $\rho_1 = \mathit{SAT} \wedge \rho_2 = \mathit{SAT}$ **then** return SAT
        if $\rho_1 = \mathit{UNSAT}$ **then** $\psi^p \leftarrow \psi^p \wedge \neg \mathit{fol2prop}(\pi_1)$
        if $\rho_2 = \mathit{UNSAT}$ **then** $\psi^p \leftarrow \psi^p \wedge \neg \mathit{fol2prop}(\pi_2)$
    return UNSAT

Delayed Theory Combination - notes

- Big improvement over Nelson–Oppen procedure
  - No modifications for underlying decision procedures.
  - Easily integrated into DPLL(T) framework.
- Disadvantage:
  - All equalities between interface variables added beforehand.
  - Possibly quadratic increase.
- Lazy implementations are possible
  - In the original paper it was because of inability of MathSAT to add new literals on-the-fly.

Model-based Theory Combination

- Goal: minimize the number of shared equalities.
  - In practice, number of local inconsistencies is much bigger than global (cross-theory) inconsistencies.
- Basic idea:
  - Each theory maintains a model for its part.
  - At certain points if two variables have the same value, a new interface equality is added to Boolean level.
  - At certain points try mutation of current model to reduce equalities.
- Example: Simplex-based decision procedure for linear arithmetic maintains assignment all the time.