### **Decision Procedures**

#### Jochen Hoenicke



Software Engineering Albert-Ludwigs-University Freiburg

### Winter Term 2019/2020

### Nelson-Oppen Theory Combination

Combining Decision Procedures: Nelson-Oppen Method

Motivation: How do we show that

 $F: 1 \leq x \land x \leq 2 \land f(x) \neq f(1) \land f(x) \neq f(2)$ 

is  $(T_E \cup T_{\mathbb{Z}})$ -unsatisfiable?

### Given

Multiple Theories  $T_i$  over signatures  $\Sigma_i$ (constants, functions, predicates) with corresponding decision procedures  $P_i$  for  $T_i$ -satisfiability.

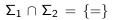
### Goal

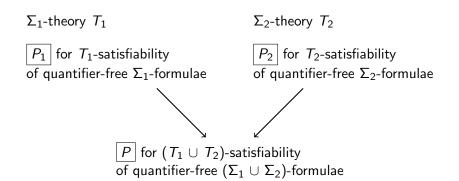
Decide satisfiability of a sentence in theory  $\cup_i T_i$ .

Jochen Hoenicke (Software Engineering)

# Nelson-Oppen Combination Method (N-O Method)







We show how to get Procedure P from Procedures  $P_1$  and  $P_2$ .

Decision Procedures

# Nelson-Oppen: Limitations

Given formula F in theory  $T_1 \cup T_2$ .

• F must be quantifier-free.

**2** Signatures  $\Sigma_i$  of the combined theory only share =, i.e.,

$$\Sigma_1 \cap \Sigma_2 = \{=\}$$

Theories must be stably infinite.

Note:

- Algorithm can be extended to combine arbitrary number of theories  $T_i$  combine two, then combine with another, and so on.
- We restrict *F* to be conjunctive formula otherwise convert to DNF and check each disjunct.



Problem: The  $T_1/T_2$ -interpretations must have the same data domain; it turns out same cardinality, e.g. infinite, is enough.

### Definition (stably infinite)

A  $\Sigma$ -theory T is stably infinite iff for every quantifier-free  $\Sigma$ -formula F: if F is T-satisfiable then there exists some infinite T-interpretation that satisfies Fwith infinite cardinality.



- $T_{\mathbb{Z}}$ : stably infinite (all *T*-interpretations are infinite).
- $T_{\mathbb{Q}}$ : stably infinite (all *T*-interpretations are infinite).
- *T*<sub>E</sub>: stably infinite (one can add infinitely many fresh and distinct values).
- Σ-theory T with Σ : {a, b, =} and axiom ∀x. x = a ∨ x = b: not stable infinite, since every T-interpretation has at most two elements.



Consider quantifier-free conjunctive ( $\Sigma_{\textit{E}} \, \cup \, \Sigma_{\mathbb{Z}})\text{-formula}$ 

$$F: 1 \leq x \wedge x \leq 2 \wedge f(x) \neq f(1) \wedge f(x) \neq f(2).$$

The signatures of  $T_E$  and  $T_{\mathbb{Z}}$  only share =. Also, both theories are stably infinite. Hence, the NO combination of the decision procedures for  $T_E$  and  $T_{\mathbb{Z}}$  decides the  $(T_E \cup T_{\mathbb{Z}})$ -satisfiability of F.

*F* is  $(T_E \cup T_{\mathbb{Z}})$ -unsatisfiable: The first two literals imply  $x = 1 \lor x = 2$  so that  $f(x) = f(1) \lor f(x) = f(2)$ . This contradicts last two literals.

# N-O Overview

# UNI

### Phase 1: Variable Abstraction

- Given conjunction  $\Gamma$  in theory  $T_1 \cup T_2$ .
- Convert to conjunction  $\Gamma_1 \cup \Gamma_2$  s.t.
  - $\Gamma_i$  in theory  $T_i$
  - $\Gamma_1 \cup \Gamma_2$  satisfiable iff  $\Gamma$  satisfiable.

### Phase 2: Check

- If there is some set S of equalities and disequalities between the shared variables of Γ<sub>1</sub> and Γ<sub>2</sub>
   shared(Γ<sub>1</sub>, Γ<sub>2</sub>) = free(Γ<sub>1</sub>) ∩ free(Γ<sub>2</sub>)
   s.t. S ∪ Γ<sub>i</sub> are T<sub>i</sub>-satisfiable for all i, then Γ is satisfiable.
- Otherwise, unsatisfiable.



Consider quantifier-free conjunctive ( $\Sigma_1 \cup \Sigma_2)\text{-formula } \textit{F}.$ 

Two versions:

- nondeterministic simple to present, but high complexity
- deterministic efficient

Nelson-Oppen (N-O) method proceeds in two steps:

- Phase 1 (variable abstraction)
  - same for both versions

### • Phase 2

nondeterministic: guess equalities/disequalities and check deterministic: generate equalities/disequalities by equality propagation



Given quantifier-free conjunctive  $(\Sigma_1 \cup \Sigma_2)$ -formula F. Transform F into two quantifier-free conjunctive formulae

 $\Sigma_1$ -formula  $F_1$  and  $\Sigma_2$ -formula  $F_2$ s.t. F is  $(T_1 \cup T_2)$ -satisfiable iff  $F_1 \wedge F_2$  is  $(T_1 \cup T_2)$ -satisfiable  $F_1$  and  $F_2$  are linked via a set of shared variables.

For term t, let hd(t) be the root symbol, e.g. hd(f(x)) = f.

## Generation of $F_1$ and $F_2$

**Generation of** 
$$F_1$$
 and  $F_2$   
For  $i, j \in \{1, 2\}$  and  $i \neq j$ , repeat the transformations  
(a) if function  $f \in \Sigma_i$  and  $hd(t) \in \Sigma_j$ ,  
 $F[f(t_1, \dots, t, \dots, t_n)]$  eqsat.  $F[f(t_1, \dots, w, \dots, t_n)] \land w = t$ 

2) if predicate 
$$p \in \Sigma_i$$
 and  $hd(t) \in \Sigma_j$ ,  
 $F[p(t_1, \ldots, t, \ldots, t_n)] \quad eqsat. \quad F[p(t_1, \ldots, w, \ldots, t_n)] \land w = t$ 

(a) if 
$$hd(s) \in \Sigma_i$$
 and  $hd(t) \in \Sigma_j$ ,  
 $F[s = t] \quad eqsat. \quad F[\top] \land w = s \land w = t$ 

• if 
$$hd(s) \in \Sigma_i$$
 and  $hd(t) \in \Sigma_j$ ,  
 $F[s \neq t] \quad eqsat. \quad F[w_1 \neq w_2] \land w_1 = s \land w_2 = t$ 

where w,  $w_1$ , and  $w_2$  are fresh variables.

Jochen Hoenicke (Software Engineering)

**Decision Procedures** 

## Example: Phase 1

Consider ( $\Sigma_E \cup \Sigma_Z$ )-formula

 $F: 1 \leq x \wedge x \leq 2 \wedge f(x) \neq f(1) \wedge f(x) \neq f(2).$ 

According to transformation 1, since  $f \in \Sigma_E$  and  $1 \in \Sigma_{\mathbb{Z}}$ , replace f(1) by  $f(w_1)$  and add  $w_1 = 1$ . Similarly, replace f(2) by  $f(w_2)$  and add  $w_2 = 2$ . Now, the literals

 $\Gamma_{\mathbb{Z}} : \{ 1 \le x, \ x \le 2, \ w_1 = 1, \ w_2 = 2 \}$ 

are  $T_{\mathbb{Z}}$ -literals, while the literals

 $\Gamma_E$ : { $f(x) \neq f(w_1), f(x) \neq f(w_2)$ }

are  $T_E$ -literals. Hence, construct the  $\Sigma_{\mathbb{Z}}$ -formula

 $F_1: 1 \leq x \land x \leq 2 \land w_1 = 1 \land w_2 = 2$ 

and the  $\Sigma_E$ -formula

$$F_2$$
:  $f(x) \neq f(w_1) \wedge f(x) \neq f(w_2)$ .

 $F_1$  and  $F_2$  share the variables  $\{x, w_1, w_2\}$ .  $F_1 \land F_2$  is  $(T_E \cup T_{\mathbb{Z}})$ -equisatisfiable to F.

## Example: Phase 1

Consider ( $\Sigma_E \cup \Sigma_{\mathbb{Z}}$ )-formula

 $F: f(x) = x + y \land x \leq y + z \land x + z \leq y \land y = 1 \land f(x) \neq f(2).$ 

In the first literal,  $hd(f(x)) = f \in \Sigma_E$  and  $hd(x + y) = + \in \Sigma_{\mathbb{Z}}$ ; thus, by (3), replace the literal with

 $w_1 = f(x) \wedge w_1 = x + y$ .

In the final literal,  $f \in \Sigma_E$  but  $2 \in \Sigma_{\mathbb{Z}}$ , so by (1), replace it with

$$f(x) \neq f(w_2) \wedge w_2 = 2$$
.

Now, separating the literals results in two formulae:

 $F_1: w_1=x+y \wedge x \leq y+z \wedge x+z \leq y \wedge y=1 \wedge w_2=2$ is a  $\Sigma_{\mathbb{Z}}$ -formula, and

$$F_2: w_1 = f(x) \wedge f(x) \neq f(w_2)$$

is a  $\Sigma_E$ -formula.

The conjunction  $F_1 \wedge F_2$  is  $(T_E \cup T_{\mathbb{Z}})$ -equisatisfiable to F.

# Phase 2: Guess and Check (Nondeterministic)

FREIBURG

- Phase 1 separated  $(\Sigma_1 \cup \Sigma_2)$ -formula F into two formulae:  $\Sigma_1$ -formula  $F_1$  and  $\Sigma_2$ -formula  $F_2$
- $F_1$  and  $F_2$  are linked by a set of shared variables:  $V = \text{shared}(F_1, F_2) = \text{free}(F_1) \cap \text{free}(F_2)$
- Let E be an equivalence relation over V.
- The arrangement  $\alpha(V, E)$  of V induced by E is:

$$\alpha(V,E) : \bigwedge_{u,v \in V.} \bigcup_{u \in v} u = v \land \bigwedge_{u,v \in V. \neg(u \in v)} u \neq v$$



#### Lemma

The original formula F is  $(T_1 \cup T_2)$ -satisfiable iff there exists an equivalence relation E of V s.t. (1)  $F_1 \wedge \alpha(V, E)$  is  $T_1$ -satisfiable, and (2)  $F_2 \wedge \alpha(V, E)$  is  $T_2$ -satisfiable.

Proof:

⇒ If F is  $(T_1 \cup T_2)$ -satisfiable, then  $F_1 \wedge F_2$  is  $(T_1 \cup T_2)$ -satisfiable, hence there is a  $T_1 \cup T_2$ -Interpretation I with  $I \models F_1 \wedge F_2$ .

Define  $E \subseteq V \times V$  with  $u \in v$  iff  $I \models u = v$ . Then E is a equivalence relation. By definition of E and  $\alpha(V, E)$ ,  $I \models \alpha(V, E)$ . Hence  $I \models F_1 \land \alpha(V, E)$  and  $I \models F_2 \land \alpha(V, E)$ . Thus, these formulae are  $T_1$ - and  $T_2$ -satisfiable, respectively.  $\leftarrow$  Let  $I_1$  and  $I_2$  be  $T_1$ - and  $T_2$ -interpretations, respectively, with

 $I_1 \models F_1 \land \alpha(V, E) \text{ and } I_2 \models F_2 \land \alpha(V, E).$ 

W.l.o.g. assume that  $\alpha_{l_1}[=](v, w)$  iff v = w iff  $\alpha_{l_2}[=](v, w)$ . (Otherwise, replace  $D_{l_i}$  with  $D_{l_i}/\alpha_{l_i}[=]$ )

Since  $T_1$  and  $T_2$  are stably infinite, we can assume that  $D_{l_1}$  and  $D_{l_2}$  are of the same cardinality.

Since 
$$I_1 \models \alpha(V, E)$$
 and  $I_2 \models \alpha(V, E)$ , for  $x, y \in V$ :  
 $\alpha_{I_1}[x] = \alpha_{I_1}[y]$  iff  $\alpha_{I_2}[x] = \alpha_{I_2}[y]$ .

Construct bijective function  $g : D_{l_1} \to D_{l_2}$  with  $g(\alpha_{l_1}[x]) = \alpha_{l_2}[x]$ for all  $x \in V$ . Define *I* as follows:  $D_I = D_{l_2}$ ,  $\alpha_I[x] = \alpha_{l_2}[x](= g(\alpha_{l_1}[x]))$  for  $x \in V$ ,  $\alpha_I[=](v,w)$  iff v = w,  $\alpha_I[f_2] = \alpha_{l_2}[f_2]$  for  $f_2 \in \Sigma_2$ ,  $\alpha_I[f_1](v_1, \ldots, v_n) = g(\alpha_{l_1}[f_1](g^{-1}(v_1), \ldots, g^{-1}(v_n)))$  for  $f_1 \in \Sigma_1$ . Then *I* is a  $T_1 \cup T_2$ -interpretation, and satisfies  $F_1 \wedge F_2$ . Hence *F* is  $T_1 \cup T_2$ -satisfiable.

Jochen Hoenicke (Software Engineering)

Decision Procedures

### Example: Phase 2

UNI FREIBURG

Consider ( $\Sigma_E \cup \Sigma_{\mathbb{Z}}$ )-formula

 $F: 1 \leq x \wedge x \leq 2 \wedge f(x) \neq f(1) \wedge f(x) \neq f(2)$ 

Phase 1 separates this formula into the  $\Sigma_{\mathbb{Z}}\text{-formula}$ 

 $F_1: 1 \leq x \land x \leq 2 \land w_1 = 1 \land w_2 = 2$ 

and the  $\Sigma_E$ -formula

$$F_2: f(x) \neq f(w_1) \land f(x) \neq f(w_2)$$

with

$$V = \text{shared}(F_1, F_2) = \{x, w_1, w_2\}$$

There are 5 equivalence relations to consider, which we list by stating the partitions:

### Example: Phase 2 (cont)

Hence, F is  $(T_E \cup T_{\mathbb{Z}})$ -unsatisfiable.

UNI FREIBURG

# Example: Phase 2 (cont)

Consider the ( $\Sigma_{\mathsf{cons}} \cup \Sigma_{\mathbb{Z}})\text{-formula}$ 

$$F : \operatorname{car}(x) + \operatorname{car}(y) = z \wedge \operatorname{cons}(x, z) \neq \operatorname{cons}(y, z) .$$

After two applications of (1), Phase 1 separates F into the  $\Sigma_{cons}$ -formula

$$F_1: w_1 = \operatorname{car}(x) \wedge w_2 = \operatorname{car}(y) \wedge \operatorname{cons}(x, z) \neq \operatorname{cons}(y, z)$$

and the  $\Sigma_{\mathbb{Z}}\text{-formula}$ 

$$F_2: w_1 + w_2 = z$$
,

with

$$V = \text{shared}(F_1, F_2) = \{z, w_1, w_2\}$$
.

Consider the equivalence relation E given by the partition

$$\{\{z\}, \{w_1\}, \{w_2\}\}$$
.

The arrangement

$$\alpha(V,E): z \neq w_1 \land z \neq w_2 \land w_1 \neq w_2$$

satisfies both  $F_1$  and  $F_2$ :  $F_1 \wedge \alpha(V, E)$  is  $T_{cons}$ -satisfiable, and  $F_2 \wedge \alpha(V, E)$  is  $T_{\mathbb{Z}}$ -satisfiable. Hence, F is  $(T_{cons} \cup T_{\mathbb{Z}})$ -satisfiable.

Jochen Hoenicke (Software Engineering)





Phase 2 was formulated as "guess and check": First, guess an equivalence relation E, then check the induced arrangement.

The number of equivalence relations grows super-exponentially with the # of shared variables. It is given by Bell numbers.

e.g., 12 shared variables  $\Rightarrow$  over four million equivalence relations.

Solution: Deterministic Version

Phase 1 as before Phase 2 asks the decision procedures  $P_1$  and  $P_2$  to propagate new equalities.

Example 1:



 $F: \quad f(f(x)-f(y)) \neq f(z) \land x \leq y \land y + z \leq x \land 0 \leq z$ 

$$F : f(f(x) - f(y)) \neq f(z) \land x \leq y \land y + z \leq x \land 0 \leq z$$

$$f(x) \Rightarrow u \quad f(y) \Rightarrow v \quad u - v \Rightarrow w$$

$$\begin{split} &\Gamma_E : \quad \{f(w) \neq f(z), \ u = f(x), \ v = f(y)\} \qquad \dots \ T_E \text{-formula} \\ &\Gamma_{\mathbb{R}} : \quad \{x \leq y, \ y + z \leq x, \ 0 \leq z, \ w = u - v\} \quad \dots \ T_{\mathbb{R}} \text{-formula} \\ & \text{shared}(\Gamma_{\mathbb{R}}, \Gamma_E) = \{x, y, z, u, v, w\} \end{split}$$

Nondeterministic version — over 200 *Es*! Let's try the deterministic version.

FREIBURG

## Phase 2: Equality Propagation

 $P_{\mathbb{R}}$  $P_E^{\perp}$  $s_0$ :  $\langle \Gamma_{\mathbb{R}}, \Gamma_F, \{\} \rangle$  $\Gamma_{\mathbb{R}} \models x = y$  $s_1$ :  $\langle \Gamma_{\mathbb{R}}, \Gamma_F, \{x = y\} \rangle$  $\Gamma_F \cup \{x = y\} \models u = v$  $s_2$ :  $\langle \Gamma_{\mathbb{R}}, \Gamma_F, \{x = y, u = v\} \rangle$  $\Gamma_{\mathbb{R}} \cup \{u = v\} \models z = w$  $s_3$ :  $\langle \Gamma_{\mathbb{R}}, \Gamma_F, \{x = y, u = v, z = w\} \rangle$  $\Gamma_F \cup \{z = w\} \models \mathsf{false}$ s₁ : false

Contradiction. Thus, F is  $(T_{\mathbb{R}} \cup T_{E})$ -unsatisfiable.

If there were no contradiction, F would be  $(T_{\mathbb{R}} \cup T_{E})$ -satisfiable.

**Decision Procedures** 

# **Convex Theories**



### Definition (convex theory)

A  $\Sigma$ -theory T is convex iff for every quantifier-free conjunction  $\Sigma$ -formula Fand for every disjunction  $\bigvee_{i=1}^{n} (u_i = v_i)$ if  $F \models \bigvee_{i=1}^{n} (u_i = v_i)$ then  $F \models u_i = v_i$ , for some  $i \in \{1, ..., n\}$ 

### Claim

Equality propagation is a decision procedure for convex theories.

Jochen Hoenicke (Software Engineering)

**Decision Procedures** 

# **Convex Theories**

- $T_E$ ,  $T_{\mathbb{R}}$ ,  $T_{\mathbb{Q}}$ ,  $T_{\text{cons}}$  are convex
- $T_{\mathbb{Z}}, T_{\mathsf{A}}$  are not convex

Example:  $T_{\mathbb{Z}}$  is not convex Consider quantifier-free conjunctive

 $F: \quad 1 \leq z \land z \leq 2 \land u = 1 \land v = 2$  $F \models z = u \lor z = v$ 

but

Then

$$\begin{array}{ccc} F & \not\models & z = u \\ F & \not\models & z = v \end{array}$$

UNI FREIBURG

#### Example:

The theory of arrays  $T_A$  is not convex. Consider the quantifier-free conjunctive  $\Sigma_A$ -formula

$$F : a\langle i \triangleleft v \rangle [j] = v$$
.

Then

$$F \Rightarrow i = j \lor a[j] = v$$
,

but

$$F \not\Rightarrow i = j$$
  
 $F \not\Rightarrow a[j] = v$ .

## What if T is Not Convex?

Case split when:

$$\Gamma \models \bigvee_{i=1}^n (u_i = v_i)$$

but

$$\Gamma \not\models u_i = v_i$$
 for all  $i = 1, \dots, n$ 

- For each i = 1, ..., n, construct a branch on which  $u_i = v_i$  is assumed.
- If all branches are contradictory, then **unsatisfiable**. Otherwise, **satisfiable**.

### Example 2: Non-Convex Theory

# $T_{\mathbb{Z}}$ not convex!



UNI FREIBURG

$$\Gamma: \left\{ \begin{array}{ll} 1 \leq x, & x \leq 2, \\ f(x) \neq f(1), & f(x) \neq f(2) \end{array} \right\} \quad \text{in } T_{\mathbb{Z}} \cup T_{E}$$

• Replace 
$$f(1)$$
 by  $f(w_1)$ , and add  $w_1 = 1$ .

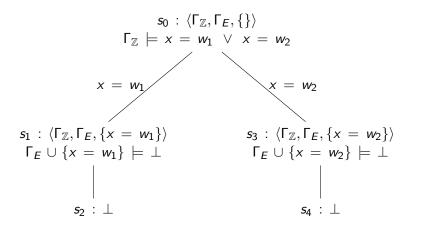
• Replace f(2) by  $f(w_2)$ , and add  $w_2 = 2$ .

Result:

$$\Gamma_{\mathbb{Z}} = \left\{ \begin{array}{l} 1 \leq x, \\ x \leq 2, \\ w_1 = 1, \\ w_2 = 2 \end{array} \right\} \quad \text{and} \quad \Gamma_E = \left\{ \begin{array}{l} f(x) \neq f(w_1), \\ f(x) \neq f(w_2) \end{array} \right\}$$

shared $(\Gamma_{\mathbb{Z}}, \Gamma_E) = \{x, w_1, w_2\}$ 

### Example 2: Non-Convex Theory



All leaves are labeled with  $\bot \Rightarrow \Gamma$  is  $(T_{\mathbb{Z}} \cup T_E)$ -unsatisfiable.

### Example 3: Non-Convex Theory

$$\Gamma: \left\{\begin{array}{cc} 1 \leq x, \ x \leq 3, \\ f(x) \neq f(1), \ f(x) \neq f(3), \ f(1) \neq f(2) \end{array}\right\} \quad \text{in } T_{\mathbb{Z}} \cup T_{E}$$

- Replace f(1) by  $f(w_1)$ , and add  $w_1 = 1$ .
- Replace f(2) by  $f(w_2)$ , and add  $w_2 = 2$ .
- Replace f(3) by  $f(w_3)$ , and add  $w_3 = 3$ .

Result:

$$\Gamma_{\mathbb{Z}} = \begin{cases} 1 \le x, \\ x \le 3, \\ w_1 = 1, \\ w_2 = 2, \\ w_3 = 3 \end{cases} \text{ and } \Gamma_E = \begin{cases} f(x) \ne f(w_1), \\ f(x) \ne f(w_3), \\ f(w_1) \ne f(w_2) \end{cases}$$
  
shared( $\Gamma_{\mathbb{Z}}, \Gamma_E$ ) = { $x, w_1, w_2, w_3$ }

UNI FREIBURG

### Example 3: Non-Convex Theory

$$s_{0} : \langle \Gamma_{\mathbb{Z}}, \Gamma_{E}, \{\} \rangle$$

$$\Gamma_{\mathbb{Z}} \models x = w_{1} \lor x = w_{2} \lor x = w_{3}$$

$$x = w_{1} \qquad x = w_{2}$$

$$s_{1} : \langle \Gamma_{\mathbb{Z}}, \Gamma_{E}, \{x = w_{1}\} \rangle s_{3} : \langle \Gamma_{\mathbb{Z}}, \Gamma_{E}, \{x = w_{2}\} \rangle s_{4} : \langle \Gamma_{\mathbb{Z}}, \Gamma_{E}, \{x = w_{3}\} \rangle$$

$$\Gamma_{E} \cup \{x = w_{1}\} \models \bot \qquad \Gamma_{E} \cup \{x = w_{3}\} \models \bot$$

$$\downarrow$$

$$s_{2} : \bot \qquad s_{5} : \bot$$

No more equations on middle leaf  $\Rightarrow \Gamma$  is  $(T_{\mathbb{Z}} \cup T_E)$ -satisfiable.

UNI FREIBURG