

Decision Procedures

Jochen Hoenicke



Software Engineering
Albert-Ludwigs-University Freiburg

Summer 2013

Nelson-Oppen Theory Combination

Motivation: How do we show that

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

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

Given

Multiple Theories T_i over signatures Σ_i
(constants, functions, predicates)

with corresponding decision procedures P_i for T_i -satisfiability.

Goal

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

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

Σ_1 -theory T_1

Σ_2 -theory T_2

P_1 for T_1 -satisfiability
of quantifier-free Σ_1 -formulae

P_2 for T_2 -satisfiability
of quantifier-free Σ_2 -formulae

P for $(T_1 \cup T_2)$ -satisfiability
of quantifier-free $(\Sigma_1 \cup \Sigma_2)$ -formulae

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

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

- 1 F must be quantifier-free.
- 2 Signatures Σ_i of the combined theory **only share** $=$, i.e.,

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

- 3 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 Σ -theory T is **stably infinite** iff
for every quantifier-free Σ -formula F :
if F is T -satisfiable
then there exists some **infinite** T -interpretation that satisfies F with **infinite cardinality**.

- $T_{\mathbb{Z}}$: stably infinite (all T -interpretations are infinite).
- $T_{\mathbb{Q}}$: stably infinite (all T -interpretations are infinite).
- T_E : stably infinite (one can add infinitely many fresh and distinct values).
- Σ -theory T with $\Sigma : \{a, b, =\}$ and axiom $\forall x. x = a \vee x = b$:
not stable infinite,
since every T -interpretation has at most two elements.

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

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

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 $x = 1 \vee x = 2$ so that $f(x) = f(1) \vee f(x) = f(2)$. This contradicts last two literals.

Phase 1: Variable Abstraction

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

Phase 2: Check

- If there is some set S of equalities and disequalities between the shared variables of Γ_1 and Γ_2
 $\text{shared}(\Gamma_1, \Gamma_2) = \text{free}(\Gamma_1) \cap \text{free}(\Gamma_2)$
s.t. $S \cup \Gamma_i$ are T_i -satisfiable for all i ,
then Γ is **satisfiable**.
- Otherwise, **unsatisfiable**.

Consider quantifier-free conjunctive $(\Sigma_1 \cup \Sigma_2)$ -formula 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

Σ_1 -formula F_1 and Σ_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 $\text{hd}(t)$ be the root symbol, e.g. $\text{hd}(f(x)) = f$.

For $i, j \in \{1, 2\}$ and $i \neq j$, repeat the transformations

- ① if function $f \in \Sigma_i$ and $\text{hd}(t) \in \Sigma_j$,

$$F[f(t_1, \dots, t, \dots, t_n)] \text{ eqsat. } F[f(t_1, \dots, w, \dots, t_n)] \wedge w = t$$

- ② if predicate $p \in \Sigma_i$ and $\text{hd}(t) \in \Sigma_j$,

$$F[p(t_1, \dots, t, \dots, t_n)] \text{ eqsat. } F[p(t_1, \dots, w, \dots, t_n)] \wedge w = t$$

- ③ if $\text{hd}(s) \in \Sigma_i$ and $\text{hd}(t) \in \Sigma_j$,

$$F[s = t] \text{ eqsat. } F[\top] \wedge w = s \wedge w = t$$

- ④ if $\text{hd}(s) \in \Sigma_i$ and $\text{hd}(t) \in \Sigma_j$,

$$F[s \neq t] \text{ eqsat. } F[w_1 \neq w_2] \wedge w_1 = s \wedge w_2 = t$$

Example: Phase 1

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

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

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

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

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

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

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

Example: Phase 1

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

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

In the first literal, $\text{hd}(f(x)) = f \in \Sigma_E$ and $\text{hd}(x + y) = + \in \Sigma_{\mathbb{Z}}$; 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 Σ_E -formula.

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

- Phase 1 **separated** $(\Sigma_1 \cup \Sigma_2)$ -formula F into two formulae:
 Σ_1 -formula F_1 and Σ_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:

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

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:

\Rightarrow 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 $u E v$ 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 $I \models F_1 \wedge \alpha(V, E)$ and $I \models F_2 \wedge \alpha(V, E)$.

Thus, these formulae are T_1 - and T_2 -satisfiable, respectively.

← Let I_1 and I_2 be T_1 - and T_2 -interpretations, respectively, with

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

W.l.o.g. assume that $\alpha_{I_1}[=](v, w)$ iff $v = w$ iff $\alpha_{I_2}[=](v, w)$.
(Otherwise, replace D_{I_i} with $D_{I_i}/\alpha_{I_i}[=]$)

Since T_1 and T_2 are stably infinite, we can assume that D_{I_1} and D_{I_2} are of the same cardinality.

Since $I_1 \models \alpha(V, E)$ and $I_2 \models \alpha(V, E)$, for $x, y \in V$:

$$\alpha_{I_1}[x] = \alpha_{I_1}[y] \text{ iff } \alpha_{I_2}[x] = \alpha_{I_2}[y].$$

Construct bijective function $g : D_{I_1} \rightarrow D_{I_2}$ with $g(\alpha_{I_1}[x]) = \alpha_{I_2}[x]$ for all $x \in V$. Define I as follows: $D_I = D_{I_2}$,

$$\alpha_I[x] = \alpha_{I_2}[x] (= g(\alpha_{I_1}[x])) \text{ for } x \in V,$$

$$\alpha_I[=](v, w) \text{ iff } v = w,$$

$$\alpha_I[f_2] = \alpha_{I_2}[f_2] \text{ for } f_2 \in \Sigma_2,$$

$$\alpha_I[f_1](v_1, \dots, v_n) = g(\alpha_{I_1}[f_1](g^{-1}(v_1), \dots, g^{-1}(v_n))) \text{ for } f_1 \in \Sigma_1.$$

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.

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 $\Sigma_{\mathbb{Z}}$ -formula

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

and the Σ_E -formula

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

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)

- ① $\{\{x, w_1, w_2\}\}$, i.e., $x = w_1 = w_2$:
 $x = w_1$ and $f(x) \neq f(w_1) \Rightarrow F_2 \wedge \alpha(V, E)$ is T_E -unsatisfiable.
- ② $\{\{x, w_1\}, \{w_2\}\}$, i.e., $x = w_1, x \neq w_2$:
 $x = w_1$ and $f(x) \neq f(w_1) \Rightarrow F_2 \wedge \alpha(V, E)$ is T_E -unsatisfiable.
- ③ $\{\{x, w_2\}, \{w_1\}\}$, i.e., $x = w_2, x \neq w_1$:
 $x = w_2$ and $f(x) \neq f(w_2) \Rightarrow F_2 \wedge \alpha(V, E)$ is T_E -unsatisfiable.
- ④ $\{\{x\}, \{w_1, w_2\}\}$, i.e., $x \neq w_1, w_1 = w_2$:
 $w_1 = w_2$ and $w_1 = 1 \wedge w_2 = 2$
 $\Rightarrow F_1 \wedge \alpha(V, E)$ is $T_{\mathbb{Z}}$ -unsatisfiable.
- ⑤ $\{\{x\}, \{w_1\}, \{w_2\}\}$, i.e., $x \neq w_1, x \neq w_2, w_1 \neq w_2$:
 $x \neq w_1 \wedge x \neq w_2$ and $x = w_1 = 1 \vee x = w_2 = 2$
 (since $1 \leq x \leq 2$ implies that $x = 1 \vee x = 2$ in $T_{\mathbb{Z}}$) $\Rightarrow F_1 \wedge \alpha(V, E)$ is $T_{\mathbb{Z}}$ -unsatisfiable.

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

Example: Phase 2 (cont)

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

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

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

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

and the $\Sigma_{\mathbb{Z}}$ -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 $\alpha(V, E) : z \neq w_1 \wedge z \neq w_2 \wedge w_1 \neq w_2$

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 $(T_{\text{cons}} \cup T_{\mathbb{Z}})$ -satisfiable.

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:

Real linear arithmetic $T_{\mathbb{R}}$

$P_{\mathbb{R}}$

Theory of equality T_E

P_E

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

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

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

$$\Gamma_E : \{f(w) \neq f(z), u = f(x), v = f(y)\} \quad \dots T_E\text{-formula}$$

$$\Gamma_{\mathbb{R}} : \{x \leq y, y + z \leq x, 0 \leq z, w = u - v\} \quad \dots T_{\mathbb{R}}\text{-formula}$$

$$\text{shared}(\Gamma_{\mathbb{R}}, \Gamma_E) = \{x, y, z, u, v, w\}$$

Nondeterministic version — over 200 E s!

Let's try the deterministic version.

$P_{\mathbb{R}}$

$$s_0 : \langle \Gamma_{\mathbb{R}}, \Gamma_E, \{\} \rangle$$

 P_E

$$\Gamma_{\mathbb{R}} \models x = y$$

$$s_1 : \langle \Gamma_{\mathbb{R}}, \Gamma_E, \{x = y\} \rangle$$

$$\Gamma_E \cup \{x = y\} \models u = v$$

$$s_2 : \langle \Gamma_{\mathbb{R}}, \Gamma_E, \{x = y, u = v\} \rangle$$

$$\Gamma_{\mathbb{R}} \cup \{u = v\} \models z = w$$

$$s_3 : \langle \Gamma_{\mathbb{R}}, \Gamma_E, \{x = y, u = v, z = w\} \rangle$$

$$\Gamma_E \cup \{z = w\} \models \text{false}$$

$$s_4 : \text{false}$$

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.

Definition (convex theory)

A Σ -theory T is **convex** iff

for every quantifier-free conjunction Σ -formula F

and 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 $i \in \{1, \dots, n\}$

Claim

Equality propagation is a decision procedure for convex theories.

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

Example: $T_{\mathbb{Z}}$ is not convex

Consider quantifier-free conjunctive

$$F : 1 \leq z \wedge z \leq 2 \wedge u = 1 \wedge v = 2$$

Then

$$F \models z = u \vee z = v$$

but

$$F \not\models z = u$$

$$F \not\models z = v$$

Example:

The theory of arrays T_A is not convex.

Consider the quantifier-free conjunctive Σ_A -formula

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

Then

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

but

$$\begin{aligned} F &\not\Rightarrow i = j \\ F &\not\Rightarrow a[j] = v . \end{aligned}$$

Case split when:

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

but

$$\Gamma \not\models u_i = v_i \quad \text{for all } i = 1, \dots, n$$

- For each $i = 1, \dots, n$, 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!

$$\boxed{P_{\mathbb{Z}}}$$

T_E convex

$$\boxed{P_E}$$

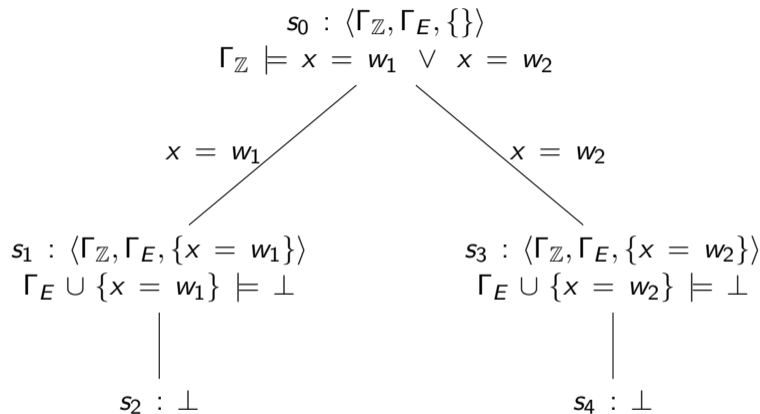
$$\Gamma : \left\{ \begin{array}{l} 1 \leq x, \quad x \leq 2, \\ f(x) \neq f(1), \quad f(x) \neq f(2) \end{array} \right\} \text{ in } T_{\mathbb{Z}} \cup T_E$$

- 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:

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

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



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

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

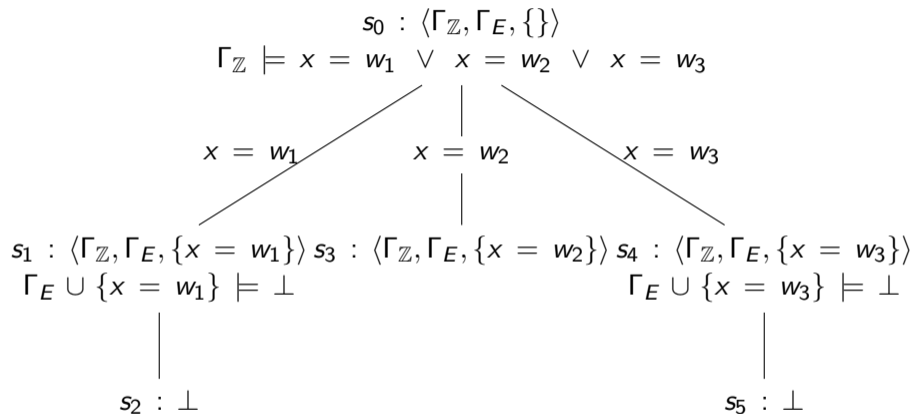
- Replace $f(1)$ by $f(w_1)$, and add $w_1 = 1$.
- Replace $f(2)$ by $f(w_2)$, and add $w_2 = 2$.
- Replace $f(3)$ by $f(w_3)$, and add $w_3 = 3$.

Result:

$$\Gamma_{\mathbb{Z}} = \left\{ \begin{array}{l} 1 \leq x, \\ x \leq 3, \\ w_1 = 1, \\ w_2 = 2, \\ w_3 = 3 \end{array} \right\} \quad \text{and} \quad \Gamma_E = \left\{ \begin{array}{l} f(x) \neq f(w_1), \\ f(x) \neq f(w_3), \\ f(w_1) \neq f(w_2) \end{array} \right\}$$

$$\text{shared}(\Gamma_{\mathbb{Z}}, \Gamma_E) = \{x, w_1, w_2, w_3\}$$

Example 3: Non-Convex Theory



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