1.8 Combining Decision Procedures

Problem:

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be first-order theories over the signatures Σ_1 and Σ_2.

Assume that we have decision procedures for the satisfiability of existentially quantified formulas (or the validity of universally quantified formulas) w. r. t. $\mathcal{T}_1$ and $\mathcal{T}_2$.

Can we combine them to get a decision procedure for the satisfiability of existentially quantified formulas w. r. t. $\mathcal{T}_1 \cup \mathcal{T}_2$?

General assumption:

Σ_1 and Σ_2 are disjoint.

The only symbol shared by $\mathcal{T}_1$ and $\mathcal{T}_2$ is built-in equality.

We consider only conjunctions of literals.

For general formulas, convert to DNF first and consider each conjunction individually.

Abstraction

To be able to use the individual decision procedures, we have to transform the original formula in such a way that each atom contains only symbols of one of the signatures (plus variables).

This process is known as *variable abstraction* or *purification*.

We apply the following rule as long as possible:

$$\frac{\exists\vec{x}\,(F[t])}{\exists\vec{x}, y\,(F[y] \land t \approx y)}$$

if the top symbol of t belongs to Σ_i and t occurs in F directly below a Σ_j-symbol or in a (positive or negative) equation $s \approx t$ where the top symbol of s belongs to Σ_j $(i \neq j)$, and if y is a new variable.

It is easy to see that the original and the purified formula are equivalent.

Stable Infiniteness

Problem:

Even if the Σ_1-formula F_1 and the Σ_2-formula F_2 do not share any symbols (not even variables), and if F_1 is $\mathcal{T}_1$-satisfiable and F_2 is $\mathcal{T}_2$-satisfiable, we cannot conclude that $F_1 \wedge F_2$ is $(\mathcal{T}_1 \cup \mathcal{T}_2)$-satisfiable.

Example:

Consider

$\mathcal{T}_1 = \{\forall x, y, z\, (x \approx y \;\vee\; x \approx z \;\vee\; y \approx z)\}$

and

$\mathcal{T}_2 = \{\exists x, y, z\, (x \not\approx y \;\wedge\; x \not\approx z \;\wedge\; y \not\approx z)\}$.

All $\mathcal{T}_1$-models have at most two elements, and all $\mathcal{T}_2$-models have at least three elements.

Since $\mathcal{T}_1 \cup \mathcal{T}_2$ is contradictory, there are no $(\mathcal{T}_1 \cup \mathcal{T}_2)$-satisfiable formulas.

To ensure that $\mathcal{T}_1$-models and $\mathcal{T}_2$-models can be combined to $(\mathcal{T}_1 \cup \mathcal{T}_2)$-models, we require that both $\mathcal{T}_1$ and $\mathcal{T}_2$ are stably infinite.

A first-order theory $\mathcal{T}$ is called *stably infinite*, if every existentially quantified formula that has a $\mathcal{T}$-model has also a $\mathcal{T}$-model with a (countably) infinite universe.

Note: By the Löwenheim–Skolem theorem, “countable” is redundant here.

Shared Variables

Even if $\exists\vec{x}\, F_1$ is $\mathcal{T}_1$-satisfiable and $\exists\vec{x}\, F_2$ is $\mathcal{T}_2$-satisfiable, it can happen that $\exists\vec{x}\, (F_1 \wedge F_2)$ is not $(\mathcal{T}_1 \cup \mathcal{T}_2)$-satisfiable, for instance because the shared variables x and y must be equal in all $\mathcal{T}_1$-models of $\exists\vec{x}\, F_1$ and different in all $\mathcal{T}_2$-models of $\exists\vec{x}\, F_2$.

Example:

Consider

$F_1 = (x + (-y) \approx 0)$,

and

$F_2 = (f(x) \not\approx f(y))$

where $\mathcal{T}_1$ is linear rational arithmetic and $\mathcal{T}_2$ is EUF.

We must exchange information about shared variables to detect the contradiction.

The Nelson–Oppen Algorithm (Non-deterministic Version)

Suppose that $\exists \vec{x}\, F$ is a purified conjunction of Σ_1 and Σ_2-literals.

Let F_1 be the conjunction of all literals of F that do not contain Σ_2-symbols; let F_2 be the conjunction of all literals of F that do not contain Σ_1-symbols. (Equations between variables are in both F_1 and F_2.)

The Nelson–Oppen algorithm starts with the pair F_1, F_2 and applies the following inference rules.

Unsat:

$$\frac{F_1, F_2}{\bot}$$

if $\exists \vec{x}\, F_i$ is unsatisfiable w. r. t. $\mathcal{T}_i$ for some i.

Branch:

$$\frac{F_1, F_2}{F_1 \wedge (x \approx y), F_2 \wedge (x \approx y) \quad | \quad F_1 \wedge (x \not\approx y), F_2 \wedge (x \not\approx y)}$$

if x and y are two different variables appearing in both F_1 and F_2 such that neither $x \approx y$ nor $x \not\approx y$ occurs in both F_1 and F_2

"|" means non-deterministic (backtracking!) branching of the derivation into two subderivations. Derivations are therefore trees. All branches need to be reduced until termination.

Clearly, all derivation paths are finite since there are only finitely many *shared variables* in F_1 and F_2, therefore the procedure represented by the rules is terminating.

We call a constraint configuration to which no rule applies *irreducible*.

Theorem 1.1 (Soundness) *If "Branch" can be applied to F_1, F_2, then $\exists \vec{x}\, (F_1 \wedge F_2)$ is satisfiable in $\mathcal{T}_1 \cup \mathcal{T}_2$ if and only if one of the successor configurations of F_1, F_2 is satisfiable in $\mathcal{T}_1 \cup \mathcal{T}_2$.*

Corollary 1.2 *If all paths in a derivation tree from F_1, F_2 end in $\bot$, then $\exists \vec{x}\, (F_1 \wedge F_2)$ is unsatisfiable in $\mathcal{T}_1 \cup \mathcal{T}_2$.*

For completeness we need to show that if one branch in a derivation terminates with an irreducible configuration F_1, F_2 (different from $\bot$), then $\exists \vec{x}\, (F_1 \wedge F_2)$ (and, thus, the initial formula of the derivation) is satisfiable in the combined theory.

As $\exists\vec{x}\,(F_1 \wedge F_2)$ is irreducible by "Unsat", the two formulas are satisfiable in their respective component theories, that is, we have $\mathcal{T}_i$-models $\mathcal{A}_i$ of $\exists\vec{x}\,F_i$ for $i \in \{1,2\}$. We are left with combining the models into a single one that is both a model of the combined theory and of the combined formula. These constructions are called *amalgamations*.

Let F be a Σ_i-formula and let S be a set of variables of F. F is called *compatible* with an equivalence $\sim$ on S if the formula

$$\exists\vec{z}\left(F \wedge \bigwedge_{x\sim y} x \approx y \wedge \bigwedge_{x,y\in S,\ x\not\sim y} x \not\approx y\right) \tag{1}$$

is $\mathcal{T}_i$-satisfiable whenever F is $\mathcal{T}_i$-satisfiable. This expresses that F does not contradict equalities between the variables in S as given by $\sim$.

Proposition 1.3 *If F_1, F_2 is a pair of conjunctions over $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively, that is irreducible by "Branch", then both F_1 and F_2 are compatible with some equivalence $\sim$ on the shared variables S of F_1 and F_2.*

Proof. If F_1, F_2 is irreducible by the branching rule, then for each pair of shared variables x and y, both F_1 and F_2 contain either $x \approx y$ or $x \not\approx y$. Choose $\sim$ to be the equivalence given by all (positive) variable equations between shared variables that are contained in F_1.

Lemma 1.4 (Amalgamation Lemma) *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two stably infinite theories over disjoint signatures Σ_1 and Σ_2. Furthermore let F_1, F_2 be a pair of conjunctions of literals over $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively, both compatible with some equivalence $\sim$ on the shared variables of F_1 and F_2. Then $F_1 \wedge F_2$ is $(\mathcal{T}_1 \cup \mathcal{T}_2)$-satisfiable if and only if each F_i is $\mathcal{T}_i$-satisfiable.*

Proof. The "only if" part is obvious.

For the "if" part, assume that each of the F_i is $\mathcal{T}_i$-satisfiable. That is, there exist models $\mathcal{A}_i$ in $\mathcal{T}_i$ and variable assigments β_i such that $\mathcal{A}_i, \beta_i \models F_i$. As the F_i are compatible with an equivalence $\sim$ on their shared variables, we may assume that the β_i also satisfy the extended conjunctions in (1) with S the set of shared variables. In particular, whenever we have two shared variables x and y, $\beta_1(x) = \beta_1(y)$ if and only if $\beta_2(x) = \beta_2(y)$. Since the theories are stably infinite we may additionally assume that the $\mathcal{A}_i$ are of cardinality ω, hence there are bijections ρ_i from the domain of $\mathcal{A}_i$ to $\mathbb{N}$ such that $\rho_1(\beta_1(x))) = \rho_2(\beta_2(x))$ for each shared variable x. Now define $\mathcal{A}$ to be the algebra having $\mathbb{N}$ as its domain; for f or P in Σ_i define $f_{\mathcal{A}}(n_1,\dots,n_k) = \rho_i(f_{\mathcal{A}_i}(\rho_i^{-1}(n_1),\dots,\rho_i^{-1}(n_k)))$ and $P_{\mathcal{A}}(n_1,\dots,n_k) \Leftrightarrow P_{\mathcal{A}_i}(\rho_i^{-1}(n_1),\dots,\rho_i^{-1}(n_k))$. Define $\beta(x) = \rho_i(\beta_i(x))$ if x is a variable occurring in F_i. By construction of the ρ_i this definition is independent of the choice of i. Clearly $\mathcal{A}|_{\Sigma_i}, \beta \models F_i$, for $i = 1,2$, hence $\mathcal{A}, \beta \models F_1 \wedge F_2$. Moreover, the reducts $\mathcal{A}|_{\Sigma_i}$ are isomorphic (via ρ_i) to $\mathcal{A}_i$ and thus are models of $\mathcal{T}_i$, so that $\mathcal{A}$ is a model of $\mathcal{T}_1 \cup \mathcal{T}_2$ as required.

Theorem 1.5 *The non-deterministic Nelson–Oppen algorithm is terminating and complete for deciding satisfiability of pure conjunctions of literals F_1 and F_2 over $\mathcal{T}_1 \cup \mathcal{T}_2$ for signature-disjoint, stably infinite theories $\mathcal{T}_1$ and $\mathcal{T}_2$.*

Proof. Suppose that F_1, F_2 is irreducible by the inference rules of the Nelson–Oppen algorithm. Applying the amalgamation lemma in combination with Prop. 1.3 we infer that F_1, F_2 is satisfiable w. r. t. $\mathcal{T}_1 \cup \mathcal{T}_2$.

Convexity

The number of possible equivalences of shared variables grows superexponentially with the number of shared variables, so enumerating all possible equivalences non-deterministically is going to be inefficient.

A much faster variant of the Nelson–Oppen algorithm exists for convex theories.

A first-order theory $\mathcal{T}$ is called *convex w. r. t. equations*, if for every conjunction Γ of Σ-equations and non-equational Σ-literals and for all Σ-equations A_i $(1 \leq i \leq n)$, whenever $\mathcal{T} \models \forall \vec{x}\,(\Gamma \to A_1 \vee \ldots \vee A_n)$, then there exists some index j such that $\mathcal{T} \models \forall \vec{x}\,(\Gamma \to A_j)$.

Theorem 1.6 *If a first-order theory $\mathcal{T}$ is convex w. r. t. equations and has non-trivial models (i. e., models with more than one element), then $\mathcal{T}$ is stably infinite.*

Proof. We shall prove the contrapositive of the statement. Suppose $\mathcal{T}$ is not stably infinite. Then there exists a satisfiable conjunction of literals $\exists \vec{x}\, F$ that has only finite models w. r. t. $\mathcal{T}$. We split F into two conjunctions F^+ and F^-, such that F^- contains the negative equational literals in F and F^+ contains the rest. As $\mathcal{T}$ is a first-order theory, it is compact, hence all models of F are bounded in cardinality by some number m. Now consider the clause $C = F^+ \to \neg F^- \vee \bigvee_{1 \leq i < j \leq m+1} y_i \approx y_j$, with fresh variables $y_1, \ldots, y_{m+1}$ not occurring in F. $\mathcal{T} \models \forall \vec{x}, \vec{y}\, C$, as the clause exactly expresses that all models of F have size less than or equal to m. However, $\mathcal{T} \not\models \forall \vec{x}, \vec{y}\,(F^+ \to A)$, for any literal A of $\neg F^-$ (as otherwise F would not be satisfiable), and also $\mathcal{T} \not\models \forall \vec{x}, \vec{y}\,(F^+ \to y_i \approx y_j)$, for each i, j, as otherwise $\mathcal{T}$ would have only trivial models, which we have excluded.

Lemma 1.7 *Suppose $\mathcal{T}$ is convex, F a conjunction of literals, and S a subset of its variables. Let, for any pair of variables x_i and x_j in S, $x_i \sim x_j$ if and only if $\mathcal{T} \models \forall\vec{x}\,(F \rightarrow x_i \approx x_j)$. Then F is compatible with $\sim$.*

Proof. We show that with this choice of $\sim$ the constraint (1) is satisfiable in $\mathcal{T}$ whenever F is. Suppose, to the contrary, that F is satisfiable but (1) is not, that is,

$$\mathcal{T} \models \forall\vec{z}\Big(F \rightarrow \bigvee_{x\sim y} x \not\approx y \;\vee \bigvee_{x,y\in S,\; x\not\sim y} x \approx y\Big)$$

or, equivalently,

$$\mathcal{T} \models \forall\vec{z}\Big(F^+ \wedge \bigwedge_{x\sim y} x \approx y \;\rightarrow\; \neg F^- \vee \bigvee_{x,y\in S,\; x\not\sim y} x \approx y\Big).$$

By convexity of $\mathcal{T}$, the antecedent implies one of the equations of the succedent. Since the equations $x \approx y$, with $x \sim y$, are entailed by F and since F is satisfiable, this means that this equation must come from the last disjunct. In other words, there exists a pair of different variables x' and y' in S such that $x' \not\sim y'$ and

$$\mathcal{T} \models \forall\vec{z}\Big(F^+ \wedge \bigwedge_{x\sim y} x \approx y \;\rightarrow\; x' \approx y'\Big).$$

Since

$$\mathcal{T} \models \forall\vec{z}\Big(F \rightarrow \bigwedge_{x\sim y} x \approx y\Big),$$

we derive $\mathcal{T} \models \forall\vec{z}\Big(F \;\rightarrow\; x' \approx y'\Big)$, which is impossible.