3.8 Integrating Theories I: E-Unification

Dealing with mathematical theories naively in a superposition prover is difficult:

Some axioms (e. g., commutativity) cannot be oriented w. r. t. a reduction ordering
⇒ provers compute many equivalent copies of a formula.

Some axiom sets (e. g., torsion-freeness, divisibility) are infinite
⇒ can we tell which axioms are really needed?

Hierarchic ("black-box") superposition is easy to implement, but conditions like compactness and sufficient completeness are rather restrictive.

Can we integrate theories directly into theorem proving calculi ("white-box" integration)?

Idea:

In order to avoid enumerating entire congruence classes w. r. t. an equational theory E, treat formulas as *representatives* of their congruence classes.

Compute an inference between formula C and D if an inference between some clause represented by C and some clause represented by D would be possible.

Consequence: We have to check whether there are substitutions that make terms s and t equal w. r. t. E
⇒ unification is replaced by E-unification.

E-Unification

E-unification (unification modulo an equational theory E):

For a set of equality problems $\{s_1 \approx t_1, \ldots, s_n \approx t_n\}$, an E-unifier is a substitution σ such that for all $i \in \{1, \ldots, n\}$: $s_i\sigma \approx_E t_i\sigma$.

Recall: $s_i\sigma \approx_E t_i\sigma$ means $E \models s_i\sigma \approx t_i\sigma$.

In general, there are infinitely many (E-)unifiers.
What about most general unifiers?

Frequent cases: $E = \emptyset$, $E = \mathrm{AC}$, $E = \mathrm{ACU}$:

$$
\begin{array}{rcll}
x + (y + z) & \approx & (x + y) + z & (\text{associativity} = \mathrm{A}) \\
x + y & \approx & y + x & (\text{commutativity} = \mathrm{C}) \\
x + 0 & \approx & x & (\text{identity (unit)} = \mathrm{U})
\end{array}
$$

The identity axiom is also abbreviated by "1", in particular, if the binary operation is denoted by $*$. (ACU = AC1).

Example:

$x + y$ and c are ACU-unifiable with $\{x \mapsto c,\ y \mapsto 0\}$ and $\{x \mapsto 0,\ y \mapsto c\}$.

$x + y$ and $x' + y'$ are ACU-unifiable with $\{x \mapsto z_1 + z_2,\ y \mapsto z_3 + z_4,\ x' \mapsto z_1 + z_3,\ y' \mapsto z_2 + z_4\}$ (among others).

More general substitutions:

Let X be a set of variables.
A substitution σ is more general modulo E than a substitution σ' on X, if there exists a substitution ρ such that $x\sigma\rho \approx_E x\sigma'$ for all $x \in X$.

Notation: $\sigma \lesssim_E^X \sigma'$.

(Why X? Because we cannot restrict to idempotent subsitutions.)

Complete sets of unifiers:

Let S be an E-unification problem, let $X = Var(S)$.
A set C of E-unifiers of S is called complete (CSU),
if for every E-unifier σ' of S there exists a $\sigma \in C$
with $\sigma \lesssim_E^X \sigma'$.

A complete set of E-unifiers C is called minimal (μCSU),
if for all $\sigma, \sigma' \in C$, $\sigma \lesssim_E^X \sigma'$ implies $\sigma = \sigma'$.

Note: every E-unification problem has a CSU. (Why?)

The set of equations E is of unification type

unitary, if every E-unification problem has a μCSU with cardinality ≤ 1 (e. g.: $E = \emptyset$);

finitary, if every E-unification problem has a finite μCSU (e. g.: $E = \mathrm{ACU}$, $E = \mathrm{AC}$, $E = \mathrm{C}$);

infinitary, if every E-unification problem has a μCSU and some E-unification problem has an infinite μCSU (e. g.: $E = \mathrm{A}$);

zero (or nullary), if some E-unification problem does not have a μCSU (e. g.: $E = \mathrm{A} \cup \{x + x \approx x\}$).

Unification modulo ACU

Let us first consider elementary ACU-unification:
the terms to be unified contain only variables and the function symbols from $\Sigma = (\{+/2,\ 0/0\}, \emptyset)$.

Since parentheses and the order of summands don't matter, every term over $X_n = \{x_1, \ldots, x_n\}$ can be written as a sum $\sum_{i=1}^{n} a_i\, x_i$.

The ACU-equivalence class of a term $t = \sum_{i=1}^{n} a_i\, x_i \in T_\Sigma(X_n)$ is uniquely determined by the vector $\vec{v}_n(t) = (a_1, \ldots, a_n)$.

Analogously, a substitution $\sigma = \{\, x_i \to \sum_{j=1}^{m} b_{ij}\, x_j \mid 1 \le i \le n \,\}$ is uniquely determined by the matrix

$$M_{n,m}(\sigma) = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}$$

Let $t = \sum_{i=1}^{n} a_i\, x_i$ and $\sigma = \{\, x_i \to \sum_{j=1}^{m} b_{ij}\, x_j \mid 1 \le i \le n \,\}$.

$$\begin{aligned} \text{Then } t\sigma &= \textstyle\sum_{i=1}^{n} a_i\, (\sum_{j=1}^{m} b_{ij}\, x_j) \\ &= \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m} a_i\, b_{ij}\, x_j \\ &= \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n} a_i\, b_{ij}\, x_j \\ &= \textstyle\sum_{j=1}^{m} (\sum_{i=1}^{n} a_i\, b_{ij})\, x_j. \end{aligned}$$

Consequence:

$$\vec{v}_m(t\sigma) = \vec{v}_n(t) \cdot M_{n,m}(\sigma).$$

Let $S = \{s_1 \approx t_1, \ldots, s_k \approx t_k\}$ be a set of equality problems over $T_\Sigma(X_n)$.

Then the following properties are equivalent:

(a) σ is an ACU-unifier of S from $X_n \to T_\Sigma(X_m)$.

(b) $\vec{v}_m(s_i\sigma) = \vec{v}_m(t_i\sigma)$ for all $i \in \{1, \ldots, k\}$.

(c) $\vec{v}_n(s_i) \cdot M_{n,m}(\sigma) = \vec{v}_n(t_i) \cdot M_{n,m}(\sigma)$ for all $i \in \{1, \ldots, k\}$.

(d) $(\vec{v}_n(s_i) - \vec{v}_n(t_i)) \cdot M_{n,m}(\sigma) = \vec{0}_m$ for all $i \in \{1, \ldots, k\}$.

(e) $M_{k,n}(S) \cdot M_{n,m}(\sigma) = \vec{0}_{k,m}$.
where $M_{k,n}(S)$ is the $k \times n$ matrix whose rows are the vectors $\vec{v}_n(s_i) - \vec{v}_n(t_i)$.

(f) The columns of $M_{n,m}(\sigma)$ are non-negative integer solutions of the system of homogeneous linear diophantine equations $DE(S)$:

$$M_{k,n}(S) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Computing unifiers:

Obviously: if $\vec{y}_1, \ldots, \vec{y}_r$ are solutions of $DE(S)$ and $a_1, \ldots, a_r$ are natural numbers, then $a_1\vec{y}_1 + \cdots + a_r\vec{y}_r$ is also a solution. (In particular, the zero vector is a solution!)

In fact, one can compute a finite set of solutions $\vec{y}_1, \ldots, \vec{y}_r$, such that every solution of $DE(S)$ can be represented as such a linear combination.

Moreover, if we combine these column vectors $\vec{y}_1, \ldots, \vec{y}_r$ to an $n \times r$ matrix, this matrix represents a most general unifier of S. (Proof: see Baader/Nipkow.)

From ACU to AC

A complete set of AC-unifiers for elementary AC-unification problems can be computed from a most general ACU-unifier by some postprocessing.

Elementary AC-unification is finitary and the elementary unifiability problem is solvable in polynomial time.

But that does not mean that minimal complete sets of AC-unifiers can be computed efficiently.

E. Domenjoud has computed the exact size of AC-μCSUs for unification problems of the following kind:

$$m\,x_1 + \cdots + m\,x_p \approx n\,y_1 + \cdots + n\,y_q$$

where $gcd(m, n) = 1$.

The number of unifiers is

$$(-1)^{p+q} \sum_{i=0}^{p} \sum_{j=0}^{q} (-1)^{i+j} \binom{p}{i} \binom{q}{j} 2^{\binom{m+j-1}{m}\binom{n+i-1}{n}}$$

For $p = m = 1$ and $q = n = 4$, that is, for the equation

$$4\,x \approx y_1 + y_2 + y_3 + y_4$$

this is

$$34\,359\,607\,481.$$

Consequence:

If possible, avoid the enumeration of AC-μCSUs
(which may have doubly exponential size).

Rather: only check AC-unifiability.

Or: use ACU instead.

Unification with Constants

So far:

Elementary unification:
terms over variables and $\{+, 0\}$ or $\{+\}$.

Step 2:

Additional free constants.

Step 3:

Additional arbitrary free function symbols.
$\rightsquigarrow$ Unification in the union of disjoint equational theories.

Unification with constants:

We can treat constants a_i like variables x_i that *must* be mapped to themselves.

Consequence: The algorithm is similar to the one we have seen before, but we have to deal with homogeneous and inhomogeneous linear diophantine equations.

Some complexity bounds change, however:

Unification type:

elementary ACU-unification: unitary;
ACU-unification with constants: finitary.

Checking unifiability:

elementary ACU-unification: trivial;
ACU-unification with constants: NP-complete.

Combining Unification Procedures

The Baader–Schulz combination procedure allows to combine unification procedures for disjoint theories (e. g., ACU and the free theory).

Basic idea (as usual): Use abstraction to convert the combined unification problem into a union of two pure unification problems; solve them individually; combine the results.

Problem 1:

The individual unification procedures might map the same variable to different terms, e. g., $\{x \mapsto y + z\}$ and $\{x \mapsto f(w)\}$.

Solution: Guess for each variable non-deterministically which procedure treats it like a constant.

Problem 2:

Combining the results might produce cycles, e. g., $\{x \mapsto y + z\}$ and $\{y \mapsto f(x)\}$.

Solution: Guess an ordering of the variables non-deterministically; each individual unifier that is computed must respect the ordering.

Note: This is a non-trivial extension that may be impossible for some unification procedures (but it is possible for *regular* equational theories, i. e., theories where for each equation $u \approx v$ the terms u and v contain the same variables).

3.9 Integrating Theories II: Calculi

We can replace syntactic unification by E-unification in the superposition calculus.

Moreover, it is usually necessary to choose a term ordering in such a way that all terms in an E-congruence class behave in the same way in comparisons (E-compatible ordering).

However, this is usually not sufficient.

AC and ACU

Example: Let $E = \mathrm{AC}$. The clauses

$$\begin{aligned} a + b &\approx d \\ b + c &\approx e \\ c + d &\not\approx a + e \end{aligned}$$

are contradictory w. r. t. AC, but if $a \succ b \succ c \succ d \succ e$, then the maximal sides of these clauses are not AC-unifiable.

We have to compute inferences if some part of a maximal sum overlaps with a part of another maximal sum (the constant b in the example above).

Technically, we can do this in such a way that we first replace positive literals $s \approx t$ by $s + x \approx t + x$, and then unify maximal sides w. r. t. AC or ACU (Peterson and Stickel 1981, Wertz 1992, Bachmair and Ganzinger 1994).

However, it turns out that even if we integrate AC or ACU in such a way into superposition, the resulting calculus is not particularly efficient – not even for ground formulas.

This is not surprising: The uniform word problem for AC or ACU is EXPSPACE-complete (Cardoza, Lipton, and Meyer 1976, Mayr and Meyer 1982).

Abelian Groups

Working in Abelian groups is easier:

If we integrate also the inverse axiom, it is sufficient to compute inferences if the maximal part of a maximal sum overlaps with the maximal part of another maximal sum (like in Gaussian elimination).

Intuitively, in Abelian groups we can always isolate the maximal part of a sum on one side of an equation.

What does that mean for the non-ground case?

Example:

$$g(y) + x \not\approx 2z \quad \vee \quad f(x) + z \approx 2y$$

Shielded variables (x, y):

occur below a free function symbol,
$\rightsquigarrow$ cannot be mapped to a maximal term,
$\rightsquigarrow$ are not involved in inferences.

Unshielded variables (z):

can be instantiated with $m \cdot u + s$, where u is maximal,
$\rightsquigarrow$ must be considered in inferences,
$\rightsquigarrow$ variable overlaps (similar to ACU).

Variable overlaps are ugly:

If we want to derive a contradiction from

$$\begin{aligned} 2a &\approx c \\ 2b &\approx d \\ 2x &\not\approx c + d \end{aligned}$$

and $a \succ b \succ c \succ d$, we have to map x to a sum of two variables $x' + x''$, unify x' with a and x'' with b.

Divible Torsion-free Abelian Groups

Working in divisible torsion-free Abelian groups is still easier:

DTAGs permit variable elimination.

Every clause can be converted into a DTAG-equivalent clause without *unshielded* variables.

Since only overlaps of maximal parts of maximal sums have to be computed, variable overlaps become unnecessary.

Moreover, if abstraction is performed eagerly, terms to be unified do not contain $+$, so ACU-unification can be replaced by standard unification.

Other Theories

A similar case: Chaining calculus for orderings.

$$\frac{D' \vee t' < t \qquad C' \vee s < s'}{(D' \vee C' \vee t' < s')\sigma}$$

where σ is a most general unifier of t and s.

Avoids explicit inferences with transitivity.
Only maximal sides of ordering literals have to be overlapped.
But unshielded variables can be maximal.

In dense linear ordering without endpoints, all unshielded variables can be eliminated.

DTAG-superposition and chaining can be combined to get a calculus for ordered divisible Abelian groups. Again, all unshielded variables can be eliminated.

Conclusion

Integrating theory axioms into superposition can become easier by integrating more axioms:

Easier unification problem (AC $\rightarrow$ ACU).

More restrictive inference rules (ACU $\rightarrow$ AG).

Less (or no) variable overlaps (AG $\rightarrow$ DTAG).

Main drawback of all theory integration methods:

For each theory, we have to start from scratch, both for the completeness proof and the implementation.