Combination of Constraint Solving Techniques: An Algebraic Point of View

Franz Baader[1] and Klaus U. Schulz[*2]

[1] Lehr- und Forschungsgebiet Theoretische Informatik, RWTH Aachen, Ahornstraße 55, 52074 Aachen, Germany, baader@informatik.rwth-aachen.de
[2] CIS, Universität München, Wagmüllerstraße 23, 80538 München, Germany, schulz@cis.uni-muenchen.de

Abstract. In a previous paper we have introduced a method that allows one to combine decision procedures for unifiability in disjoint equational theories. Lately, it has turned out that the prerequisite for this method to apply—namely that unification with so-called linear constant restrictions is decidable in the single theories—is equivalent to requiring decidability of the positive fragment of the first order theory of the equational theories. Thus, the combination method can also be seen as a tool for combining decision procedures for positive theories of free algebras defined by equational theories. Complementing this logical point of view, the present paper isolates an abstract algebraic property of free algebras—called combinability—that clarifies why our combination method applies to such algebras. We use this algebraic point of view to introduce a new proof method that depends on abstract notions and results from universal algebra, as opposed to technical manipulations of terms (such as ordered rewriting, abstraction functions, etc.) With this proof method, the previous combination results for unification can easily be extended to the case of constraint solvers that also take relational constraints (such as ordering constraints) into account. Background information from universal algebra about free structures is given to clarify the algebraic meaning of our results.

1 Introduction

In most of the applications of unification modulo an equational theory E, a unification algorithm for elementary E-unification (which treats just the terms built over the signature of E) is not sufficient. Usually, there are at least additional free function symbols present, or even symbols defined by another equational theory. For this reason, the *combination problem* for unification algorithms is an important research topic in unification theory. Informally, this problem can be described as follows: Let E and F be equational theories over disjoint signatures, and assume that unification algorithms for E and for F are given. How can we combine these algorithms to obtain a unification algorithm for $E \cup F$. Originally,

[*] This author was supported by the EC Working Group CCL, EP6028.

the term "unification algorithm" referred to an algorithm that computes a complete set of unifiers (see [SS89, Bou93] for the most recent results on combining such algorithms). With the development of constraint approaches to theorem proving [Bür91, NiR94] and term rewriting [KK89], the role of algorithms that compute complete sets of unifiers is more and more taken on by algorithms that decide solvability of the unification problems. In this setting, more general constraints than the equational constraints $s = t$ of unification problems become important as well. For example, one might be interested in ordering constraints of the form $s \leq t$ on terms [CT94], where the predicate $\leq$ could be interpreted as the subterm ordering or as a reduction ordering.

For unification, the problem of combining decision procedures has been solved in [BS92] in a rather general way. The main tool of this combination method is a *decomposition algorithm*, which separates a given unification problem Γ of the joined theory (i.e., an $(E \cup F)$-unification problem) into pure unification subproblems Γ_E and Γ_F of the single theories. Solutions of these pure problems must satisfy additional conditions, called *linear constant restrictions* in [BS92], to yield a solution of Γ. The main result of [BS92] is that solvability of unification problems in the combined theory $E \cup F$ is decidable, provided that solvability of unification problems with linear constant restrictions is decidable in E and F. It should be noted that this result can easily be lifted to solvability of $(E \cup F)$-unification problems with linear constant restrictions. This combination result has been generalized to disunification [BS93a] and to unification in the union of theories with shared constant symbols [Rin92]. In both cases, the decomposition algorithm of [BS92] could be adapted to the new problem without serious modifications. An important goal of the present paper is to give an abstract characterization of the situations in which this seemingly ubiquitous decomposition method can be applied. In addition to the better understanding of the underlying principles, this could also yield a better basis for further generalizations. The proof method used in [BS92, BS93a, Rin92]—which depends on an infinite ordered rewrite system obtained by unfailing completion, term abstraction functions, etc.—seems not to facilitate such an abstract view (see, e.g., the rather technical "shared constructor" condition in [DKR94]).

At first sight, the notion of "unification with linear constant restrictions" is just a technical notion that makes our combination machinery work, but seems to have little further significance. In [BS93] it is shown, however, that E-unification with linear constant restrictions is decidable iff the positive fragment of the first-order theory of E is decidable. Since the positive theory of E coincides with the positive theory of the E-free Σ-algebra $\mathcal{T}(\Sigma, X)/{=_E}$ over infinitely many generators X, the combination result of [BS92] can be reformulated as follows: Let E and F be equational theories over disjoint signatures Σ and Δ, and let X be a countably infinite set of generators. The positive theory of $\mathcal{T}(\Sigma \cup \Delta, X)/{=_{E\cup F}}$ is decidable, provided that the positive theories of $\mathcal{T}(\Sigma, X)/{=_E}$ and $\mathcal{T}(\Delta, X)/{=_F}$ are decidable.

In the present paper, this observation is used as the starting point of a more abstract, algebraic approach to formulating and solving the combination problem. Starting with two algebras over disjoint signatures, the goal is to construct a "combined" algebra such that validity of positive formulae in this algebra can be decided by using a decomposition algorithm and decision procedures for the positive theories of the original algebras. Obviously, this can only be achieved if the algebras satisfy some additional properties. *We will call an algebra $\mathcal{A}$* combinable *iff it is generated by a countably infinite set X such that any mapping from a finite subset of X to $\mathcal{A}$ can be extended to a surjective endomorphism of $\mathcal{A}$.* For combinable algebras $\mathcal{A}$ and $\mathcal{B}$ over disjoint signatures Σ and Δ, we can construct the so-called *free amalgamated product* $\mathcal{A} \odot \mathcal{B}$, which is a $(\Sigma \cup \Delta)$-algebra.[3] Now a simple modification of the decomposition algorithm of [BS92] can be used to show that the positive theory of $\mathcal{A} \odot \mathcal{B}$ is decidable iff the positive theories of $\mathcal{A}$ and of $\mathcal{B}$ are decidable.

Obviously, the free algebras $\mathcal{T}(\Sigma, X)/{=_E}$ and $\mathcal{T}(\Delta, X)/{=_F}$ over a countably infinite set of generators X are combinable. In this case, the free amalgamated product yields an algebra that is isomorphic to the combined free algebra $\mathcal{T}(\Sigma \cup \Delta, X)/{=_{E \cup F}}$. Thus, the combination result of [BS92] is obtained as a corollary. As described until now, the amalgamation of combinable algebras does not yield a real generalization of this result. Indeed, one can use well-known results from universal algebra to show that an algebra is combinable (as defined above) iff it is a free algebra over countably many generators for an equational theory. What is new, though, is the proof method, which—in contrast to the original proof—only depends on elementary notions from universal algebra, and thus clarifies the role played by the combinability condition. This new proof can be seen as an adaptation of the proof ideas in [SS89] to the combination of decision procedures. Unlike in [SS89], however, everything is done on the abstract algebraic level instead of on the term level. Interestingly, on this level it is also very easy to prove completeness of an optimized version of the decomposition algorithm of [BS92], which significantly reduces the number of nondeterministic choices.

In addition, the abstract algebraic approach allows for an easier generalization of the results. In fact, instead of algebras we will consider algebraic structures in the following. This means that the signatures may contain both function symbols and predicate symbols, and these additional predicate symbols may occur in the constraint problems to be solved. With the usual notion of homomorphism for structures, most of the results from universal algebra carry over to structures. The combination result for combinable algebras sketched above thus holds for free structures as well. This yields a combination method for constraint solvers of more general constraints than just equational constraints.

The next section recalls some results from universal algebra for free structures. In Section 3 we construct the free amalgamated product of free structures, and show that it again yields a free structure. Section 4 describes the decomposition algorithm and proves that it is sound and complete for existential positive input formulae. In Subsection 4.3, this result is extended to positive formulae with arbitrary quantifier prefix.

[3] This construction is similar to the one made in [SS89] for free algebras.

2 Free Structures

Let Σ be a signature consisting of a finite set Σ_F of function symbols and a finite set Σ_P of predicate symbols, where each symbol has a fixed arity. We assume that equality $=$ is an additional predicate symbol that does not occur in Σ_P. An *atomic Σ-formula* is an *equation* $s = t$ between Σ_F-terms s, t, or a *relational atomic formula* of the form $p[s_1, \ldots, s_m]$ where p is a predicate symbol in Σ_P of arity m and $s_1, \ldots, s_m$ are Σ_F-terms. A positive Σ-matrix is any Σ-formula obtained from atomic Σ-formulae using conjunction and disjunction only. A positive Σ-formula is obtained from a positive Σ-matrix by adding an arbitrary quantifier prefix, and an existential positive Σ-formula is a positive formula where the prefix consists of existential quantifier only. As usual, we shall sometimes write $t(v_1, \ldots, v_n)$ (resp. $\varphi(v_1, \ldots, v_n)$) to express that t (resp. φ) is a term (resp. formula) whose (free) variables are a subset of $\{v_1, \ldots, v_n\}$. Sentences are formulae without free variables.

A *Σ-structure* $\mathcal{A}$ has a non-empty carrier set A, and it interprets each $f \in \Sigma_F$ of arity n as an n-ary function $f_\mathcal{A}$ on A, and each $p \in \Sigma_P$ of arity m as an m-ary relation $p_\mathcal{A}$. For a formula $\varphi = \varphi(v_1, \ldots, v_n)$, we write $\mathcal{A} \models \varphi(a_1, \ldots, a_n)$ to express that φ is true in $\mathcal{A}$ under the evaluation $\{v_1 \mapsto a_1, \ldots, v_n \mapsto a_n\}$.

Usually, *Σ-constraints* are formulae of the form $\varphi(v_1, \ldots, v_n)$ with free variables. A solution of such a constraint (in a fixed Σ-structure $\mathcal{A}$) is an evaluation $\{v_1 \mapsto a_1, \ldots, v_n \mapsto a_n\}$ such that $\mathcal{A} \models \varphi(a_1, \ldots, a_n)$. Obviously, the constraint $\varphi(v_1, \ldots, v_n)$ has a solution in $\mathcal{A}$ iff the formula $\exists v_1 \ldots \exists v_n\ \varphi(v_1, \ldots, v_n)$ is valid in $\mathcal{A}$. In the present paper, we are only interested in solvability of constraints, and will thus usually take this logical point of view. In the following, tuples of variables will often be abbreviated by $\vec{v}, \vec{u}, \vec{w}$, and tuples of elements of a structure by $\vec{a}, \vec{b}$, etc. Substructures and direct products of structures are defined in the usual way. If the Σ-substructure $\mathcal{B}$ of $\mathcal{A}$ is generated by $X \subseteq A$, we write $\mathcal{B} = \langle X \rangle_\Sigma$. Later on, we will consider several signatures simultaneously. If Δ is a subset of the signature Σ, then any Σ-structure $\mathcal{A}$ can be considered as a Δ-structure (called the Δ-reduct of $\mathcal{A}$) by just forgetting about the interpretation of the additional symbols. To make clear with respect to which signature a given Σ-structure $\mathcal{A}$ is currently considered we will sometimes write $\mathcal{A}^\Sigma$ for the full Σ-structure and $\mathcal{A}^\Delta$ for its Δ-reduct.

A Σ-homomorphism is a mapping h between Σ-structures $\mathcal{A}$ and $\mathcal{B}$ such that

$$\begin{aligned} h(f_\mathcal{A}(a_1, \ldots, a_n)) &= f_\mathcal{B}(h(a_1), \ldots, h(a_n)) \\ p_\mathcal{A}[a_1, \ldots, a_n] &\Rightarrow p_\mathcal{B}[h(a_1), \ldots, h(a_n)] \end{aligned}$$

for all $f \in \Sigma_F$, $p \in \Sigma_P$, $a_1, \ldots, a_n \in A$. A Σ-isomorphism is a bijective Σ-homomorphism whose inverse is also a Σ-homomorphism.

There is an interesting (well-known) connection between surjective homomorphisms and positive formulae. The following lemma (see [Mal73], pp. 143, 144, for a proof), and its relationship to the concept of combinability, turns out to be crucial for the new proof method introduced in Section 4.

Lemma 1. *Let $h : \mathcal{A} \to \mathcal{B}$ be a surjective homomorphism between the Σ-structures $\mathcal{A}$ and $\mathcal{B}$, $\varphi(v_1, \ldots, v_m)$ be a positive Σ-formula, and $a_1, \ldots, a_m$ be elements of A. Then $\mathcal{A} \models \varphi(a_1, \ldots, a_m)$ implies $\mathcal{B} \models \varphi(h(a_1), \ldots, h(a_m))$.*

As for the case of algebras, Σ-*varieties* are defined as classes of Σ-structures that are closed under direct products, substructures, and homomorphic images. The well-known Birkhoff Theorem says that a class of Σ_F-algebras is a variety iff it is an equational class, i.e., the class of models of a set of equations. For structures, a similar characterization is possible [Mal71]: A class $\mathcal{V}$ of Σ-structures is a Σ-variety if, and only if, there exists a set E of atomic Σ-formulae[4] such that $\mathcal{V}$ is the class of models of E. In this situation, we say that $\mathcal{V}$ is the Σ-variety defined by E, and we write $\mathcal{V} = \mathcal{V}(E)$.

As in the case of varieties of algebras, varieties of structures always have free objects. Recall that a Σ-structure $\mathcal{A}$ is *free for the class of Σ-structures $\mathcal{K}$ over the set X* iff *(1)* $\mathcal{A} \in \mathcal{K}$, *(2)* $\mathcal{A}$ is generated by X, and *(3)* every mapping from X into the carrier of a Σ-structure $\mathcal{B} \in \mathcal{K}$ can be extended to a Σ-homomorphism of $\mathcal{A}$ into $\mathcal{B}$.

If $\mathcal{A}$ and $\mathcal{B}$ are free Σ-structures for the same class $\mathcal{K}$, and if their sets of generators have the same cardinality then these structures are isomorphic. Every non-trivial variety contains free structures with sets of generators of arbitrary cardinality [Mal71]. Conversely, free structures are always free for some variety [Mal71, Coh65].

Theorem 2. *Let $\mathcal{A}$ be a Σ-structure that is generated by X. Then $\mathcal{A}$ is free over X for $\{\mathcal{A}\}$ iff $\mathcal{A}$ is free over X for some Σ-variety.*

In the following, a Σ-structure $\mathcal{A}$ will be called free (over X) iff it is free (over X) for $\{\mathcal{A}\}$. Let us now analyze how free Σ-structures look like (see [Mal71, Wea93] for more information). Obviously, the Σ_F-reduct of such a structure is a free Σ_F-algebra, and thus it is (isomorphic to) an E-free Σ_F-algebra $\mathcal{T}(\Sigma_F, X)/{=_E}$ for an equational theory E. In particular, the $=_E$-equivalence classes $[s]$ of Σ_F-terms constitute the carrier of $\mathcal{A}$. It remains to be shown how the predicate symbols are interpreted on this carrier. Since $\mathcal{A}$ is free over X, any mapping from X into $T(\Sigma_F, X)/{=_E}$ can be extended to a Σ-endomorphism of $\mathcal{A}$. This, together with the definition of homomorphisms of structures, shows that the interpretation of the predicates must be closed under substitution, i.e., for all $p \in \Sigma_P$, all substitutions σ, and all terms $s_1, \ldots, s_m$, if $p[[s_1], \ldots, [s_m]]$ holds in $\mathcal{A}$ then $p[[s_1\sigma], \ldots, [s_m\sigma]]$ must also hold in $\mathcal{A}$. Conversely, it is easy to see that any extension of the Σ_F-algebra $\mathcal{T}(\Sigma_F, X)/{=_E}$ to a Σ-structure that satisfies this property is a free Σ-structure over X.

Example 1. Let Σ_F be an arbitrary set of function symbols, and assume that Σ_P consists of a single binary predicate symbol $\leq$. Consider the (absolutely free) term algebra $\mathcal{T}(\Sigma_F, X)$. We can extend this algebra to a Σ-structure by

[4] As usual, open formulae are here considered as implicitly universally quantified.

interpreting $\leq$ as subterm ordering. Another possibility would be to take a reduction ordering [Der87] such as the lexicographic path ordering. In both cases, we have closure under substitution, which means that we obtain a free Σ-structure.

Free structures over countably infinite sets of generators are canonical for the positive theory of their variety in the following sense:

Theorem 3. *Let $\mathcal{A}$ be free over the countably infinite set X for a Σ-variety $\mathcal{V}(E)$, and let ϕ be a positive Σ-formula. Then ϕ is valid in all elements of $\mathcal{V}(E)$ (i.e., ϕ is a logical consequence of E) iff ϕ is valid in $\mathcal{A}$.*

For the purpose of this paper, the following characterization of free structures is useful (see [BS94] for a proof).

Lemma 4. *Let $\mathcal{A}$ be a Σ-structure that is generated by the countably infinite set X. Then the following conditions are equivalent:*

1. *$\mathcal{A}$ is free over X.*
2. *For every finite subset X_0 of X, every mapping $h_0 : X_0 \longrightarrow A$ can be extended to a* surjective *endomorphism of $\mathcal{A}$.*

Condition 2 is the *combinability condition* mentioned in the introduction. Thus Theorem 2 shows that *a structure is free for some variety iff it is combinable.* This observation, together with Lemma 1, can be used to obtain the following lemma, which is the key tool in the new proof of correctness of the combination method.

Lemma 5. *Let $\mathcal{A}$ be a free Σ-structure over the countably infinite set of generators X, and let $\gamma = \forall\vec{u}_1\exists\vec{v}_1 \ldots \forall\vec{u}_k\exists\vec{v}_k\ \varphi(\vec{u}_1, \vec{v}_1, \ldots, \vec{u}_k, \vec{v}_k)$ be a positive Σ-sentence. Then the following conditions are equivalent:*

1. *$\mathcal{A} \models \forall\vec{u}_1\exists\vec{v}_1 \ldots \forall\vec{u}_k\exists\vec{v}_k\ \varphi(\vec{u}_1, \vec{v}_1, \ldots, \vec{u}_k, \vec{v}_k)$.*
2. *There exist tuples $\vec{x}_1 \in \vec{X}, \vec{e}_1 \in \vec{A}, \ldots, \vec{x}_k \in \vec{X}, \vec{e}_k \in \vec{A}$ and finite subsets $Z_1, \ldots, Z_k$ of X such that*
   (a) *$\mathcal{A} \models \varphi(\vec{x}_1, \vec{e}_1, \ldots, \vec{x}_k, \vec{e}_k)$,*
   (b) *all generators occurring in the tuples $\vec{x}_1, \ldots, \vec{x}_k$ are distinct,*
   (c) *for all $j, 1 \leq j \leq k$, the components of $\vec{e}_j$ are generated by Z_j, i.e., are elements of $\langle Z_j \rangle_\Sigma$, and*
   (d) *for all $j, 1 < j \leq k$, no component of $\vec{x}_j$ occurs in $Z_1 \cup \ldots \cup Z_{j-1}$.*

As an example, assume that $\mathcal{A}$ is the (absolutely free) term algebra $\mathcal{T}(\{g\}, X)$ over a signature consisting of a unary symbol g. The formula $\exists v \forall u\ g(u) = g(v)$ is not valid. In fact, Condition 2 is not satisfied since for all $x \in X$ and $e \in \mathcal{T}(\{g\}, X)$, $g(x) = g(e)$ is only satisfied in $\mathcal{T}(\{g\}, X)$ if $e = x$, and thus x is contained in any generating set of e.

Readers who are familiar with the notion of "unification with linear constant restrictions" should note the close connection between (c) and (d) of Condition 2 (which say that a generator $x \in X$ must not be contained in the generating set Z for $e \in A$ if e comes before x in the sequence $\vec{x}_1, \vec{e}_1, \ldots, \vec{x}_k, \vec{e}_k$) and a linear constant restriction (which says that the constant a must not occur in the image e of a variable u if u comes before a in the linear constant restriction).

3 Amalgamation of Free Structures

Let Σ and Δ be disjoint signatures, and let X be a countably infinite set (of generators). Let $\mathcal{A}$ be a free Σ-structure over X and and let $\mathcal{B}$ be a free Δ-structure over X. Equivalently, $\mathcal{A}$ is free over X for some Σ-variety $\mathcal{V}(E)$ and $\mathcal{B}$ is free over X for some Δ-variety $\mathcal{V}(F)$ (by Theorem 2). The following construction yields a $(\Sigma \cup \Delta)$-structure $\mathcal{A} \odot \mathcal{B}$ that is free over X for the $(\Sigma \cup \Delta)$-variety $\mathcal{V}(E \cup F)$.

We consider two countably infinite supersets X_∞ and Y_∞ of $X_0 := Y_0 := X$ such that $X_\infty \cap Y_\infty = X$ and $X_\infty \setminus X_0$ and $Y_\infty \setminus Y_0$ are infinite. Let $\mathcal{A}_\infty$ be free for $\mathcal{V}(E)$ over X_∞, and let $\mathcal{B}_\infty$ be free for $\mathcal{V}(F)$ over Y_∞. Obviously, $\mathcal{A}$ is the substructure of $\mathcal{A}_\infty$ that is generated by $X_0 \subseteq X_\infty$. Since both structures are free for the same variety, and since their generating sets X_0 and X_∞ have the same cardinality, $\mathcal{A}$ and $\mathcal{A}_\infty$ are isomorphic. The same holds for $\mathcal{B}$ and $\mathcal{B}_\infty$.

We shall make a zig-zag construction that defines ascending towers of Σ-structures $\mathcal{A}_n$ and Δ-structures $\mathcal{B}_n$. These structures are connected by bijective mappings h_n and g_n. The free amalgamated product $\mathcal{A} \odot \mathcal{B}$ will be obtained as the limit structure, which obtains its functional and relational structure from both towers by means of the limits of the mappings h_n and g_n.

$n = 0$: Let $\mathcal{A}_0 := \mathcal{A} = \langle X_0 \rangle_\Sigma$. We interpret the "new" elements in $A_0 \setminus X_0$ as generators in $\mathcal{B}_\infty$. For this purpose, select a subset $Y_1 \subseteq Y_\infty$ such that $Y_1 \cap Y_0 = \emptyset$, $|Y_1| = |A_0 \setminus X_0|$, and the remaining complement $Y_\infty \setminus (Y_0 \cup Y_1)$ is countably infinite. Choose any bijection $h_0 : Y_0 \cup Y_1 \to A_0$ where $h_0|_{Y_0} = id_{Y_0}$.

Let $\mathcal{B}_0 := \langle Y_0 \rangle_\Delta$. As for $\mathcal{A}_0$, we interpret the "new" elements in $B_0 \setminus Y_0$ as generators in $\mathcal{A}_\infty$. Select a subset $X_1 \subseteq X_\infty$ such that $X_1 \cap X_0 = \emptyset$, $|X_1| = |B_0 \setminus Y_0|$ and the remaining complement $X_\infty \setminus (X_0 \cup X_1)$ is countably infinite. Choose any bijection $g_0 : X_0 \cup X_1 \to B_0$ where $g_0|_{X_0} = id_{X_0}$.

$n \to n+1$: Suppose that $\mathcal{A}_n = \langle \bigcup_{i=0}^{n} X_i \rangle_\Sigma$ and $\mathcal{B}_n = \langle \bigcup_{i=0}^{n} Y_i \rangle_\Delta$ are already defined, and that subsets X_{n+1} of X_∞ and Y_{n+1} of Y_∞ are already given. We assume that the complements $X_\infty \setminus \bigcup_{i=0}^{n+1} X_i$ and $Y_\infty \setminus \bigcup_{i=0}^{n+1} Y_i$ are infinite, and that the sets X_i (resp. Y_i) are pairwise disjoint. In addition, we assume that bijections $h_n : B_{n-1} \cup Y_n \cup Y_{n+1} \to A_n$ and $g_n : A_{n-1} \cup X_n \cup X_{n+1} \to B_n$ are defined such that

$$(*) \quad g_n(h_n(b)) = b \text{ for } b \in B_{n-1} \cup Y_n \text{ and } h_n(g_n(a)) = a \text{ for } a \in A_{n-1} \cup X_n$$
$$(**) \quad h_n(Y_{n+1}) = A_n \setminus (A_{n-1} \cup X_n) \quad \text{and } g_n(X_{n+1}) = B_n \setminus (B_{n-1} \cup Y_n).$$

Note that $(**)$ implies that $h_n(B_{n-1} \cup Y_n) = A_{n-1} \cup X_n$ and $g_n(A_{n-1} \cup X_n) = B_{n-1} \cup Y_n$.

We define $\mathcal{A}_{n+1} = \langle \bigcup_{i=0}^{n+1} X_i \rangle_\Sigma$ and $\mathcal{B}_{n+1} = \langle \bigcup_{i=0}^{n+1} Y_i \rangle_\Delta$, and select subsets $Y_{n+2} \subseteq Y_\infty$ and $X_{n+2} \subseteq X_\infty$ such that $Y_{n+2} \cap \bigcup_{i=0}^{n+1} Y_i = \emptyset = X_{n+2} \cap \bigcup_{i=0}^{n+1} X_i$. In addition, the cardinalities must satisfy $|Y_{n+2}| = |A_{n+1} \setminus (A_n \cup X_{n+1})|$ and $|X_{n+2}| = |B_{n+1} \setminus (B_n \cup Y_{n+1})|$, and the remaining complements $Y_\infty \setminus \bigcup_{i=0}^{n+2} Y_i$ and $X_\infty \setminus \bigcup_{i=0}^{n+2} X_i$ must be countably infinite. Let

$$\upsilon_{n+1} : Y_{n+2} \to A_{n+1} \setminus (A_n \cup X_{n+1}) \quad \text{and} \quad \xi_{n+1} : X_{n+2} \to B_{n+1} \setminus (B_n \cup Y_{n+1})$$

be arbitrary bijections. We define $h_{n+1} := \upsilon_{n+1} \cup g_n^{-1} \cup h_n$ and $g_{n+1} := \xi_{n+1} \cup$

$h_n^{-1} \cup g_n$. In more detail:

$$h_{n+1}(b) = \begin{cases} \upsilon_{n+1}(b) & \text{for } b \in Y_{n+2} \\ h_n(b) & \text{for } b \in B_{n-1} \cup Y_n \cup Y_{n+1} \\ g_n^{-1}(b) & \text{for } b \in B_n \setminus (B_{n-1} \cup Y_n) \end{cases}$$

and

$$g_{n+1}(a) = \begin{cases} \xi_{n+1}(a) & \text{for } a \in X_{n+2} \\ g_n(a) & \text{for } a \in A_{n-1} \cup X_n \cup X_{n+1} \\ h_n^{-1}(a) & \text{for } a \in A_n \setminus (A_{n-1} \cup X_n). \end{cases}$$

Without loss of generality we may assume (for notational convenience) that the construction eventually covers all generators in X_∞ and Y_∞; in other words, we assume that $\bigcup_{i=0}^{\infty} X_i = X_\infty$ and $\bigcup_{i=0}^{\infty} Y_i = Y_\infty$, and thus $\bigcup_{i=0}^{\infty} A_i = A_\infty$ and $\bigcup_{i=0}^{\infty} B_i = B_\infty$. We define the limit mappings

$$h_\infty := \bigcup_{i=0}^{\infty} h_i : B_\infty \to A_\infty, \quad \text{and} \quad g_\infty := \bigcup_{i=0}^{\infty} g_i : A_\infty \to B_\infty.$$

It is easy to see that h_∞ and g_∞ are bijections that are inverse to each other. They may be used to carry the Δ-structure of $\mathcal{B}_\infty$ to $\mathcal{A}_\infty$ and to carry the Σ-structure of $\mathcal{A}_\infty$ to $\mathcal{B}_\infty$: let f (f') be an n-ary function symbol of Δ (Σ), let p (p') be an n-ary predicate symbol of Δ (Σ), and $a_1, \ldots, a_n \in A_\infty$ ($b_1, \ldots, b_n \in B_\infty$). We define

$$\begin{aligned} f_{\mathcal{A}_\infty}(a_1, \ldots, a_n) &:= h_\infty(f_{\mathcal{B}_\infty}(g_\infty(a_1), \ldots, g_\infty(a_n))), \\ f'_{\mathcal{B}_\infty}(b_1, \ldots, b_n) &:= g_\infty(f'_{\mathcal{A}_\infty}(h_\infty(b_1), \ldots, h_\infty(b_n))), \\ p_{\mathcal{A}_\infty}[a_1, \ldots, a_n] &:\Longleftrightarrow p_{\mathcal{B}_\infty}[g_\infty(a_1), \ldots, g_\infty(a_n)], \\ p'_{\mathcal{B}_\infty}[b_1, \ldots, b_n] &:\Longleftrightarrow p'_{\mathcal{A}_\infty}[h_\infty(b_1), \ldots, h_\infty(b_n)]. \end{aligned}$$

With this definition, the mappings h_∞ and g_∞ are inverse isomorphisms between the $(\Sigma \cup \Delta)$-structures $\mathcal{A}_\infty$ and $\mathcal{B}_\infty$. Identifying isomorphic structures, we call $\mathcal{A}_\infty^{\Sigma \cup \Delta} \simeq \mathcal{B}_\infty^{\Sigma \cup \Delta}$ the *free amalgamated product* $\mathcal{A} \odot \mathcal{B}$ *of* $\mathcal{A}$ *and* $\mathcal{B}$. The layering of the domain A_∞ (resp. B_∞) of this structure into the sets A_n (resp. B_n) will become important in the proof of our combination result. As a Σ-structure, $\mathcal{A} \odot \mathcal{B}$ is isomorphic to $\mathcal{A}$, which is free over X for $\mathcal{V}(E)$, and as a Δ-structure it is isomorphic to $\mathcal{B}$, which is free over X for $\mathcal{V}(F)$. In [BS94] it is shown that as a $(\Sigma \cup \Delta)$-structure it is free over X for $\mathcal{V}(E \cup F)$.

Theorem 6. *Let Σ and Δ be disjoint signatures, and let $\mathcal{A}$ be free over X for the Σ-variety $\mathcal{V}(E)$ and $\mathcal{B}$ be free over X for the Δ-variety $\mathcal{V}(F)$, where X is countably infinite. Then $\mathcal{A} \odot \mathcal{B}$ is free over X for the $(\Sigma \cup \Delta)$-variety $\mathcal{V}(E \cup F)$.*

4 Combination results

As in the previous section, let $\mathcal{V}(E)$ be a Σ-variety and $\mathcal{V}(F)$ be a Δ-variety, where Σ and Δ are disjoint signatures. For a countably infinite set of generators X, let $\mathcal{A}$ be free for $\mathcal{V}(E)$ over X, and let $\mathcal{B}$ be free for $\mathcal{V}(F)$ over X. We know that the positive theories of $\mathcal{V}(E)$ and $\mathcal{A}$ (resp. $\mathcal{V}(F)$ and $\mathcal{B}$) coincide (by Theorem 3), and that the free amalgamated product $\mathcal{A} \odot \mathcal{B}$ is free for $\mathcal{V}(E \cup F)$ over X (by Theorem 6).

In the first part of this section, we consider only existential positive $(\Sigma \cup \Delta)$-sentences. The decomposition algorithm described below can be used to reduce validity of such sentences in $\mathcal{A} \odot \mathcal{B}$ (or, equivalently, in $\mathcal{V}(E \cup F)$) to validity of positive sentences in $\mathcal{A}$ and in $\mathcal{B}$. At the end of the section we shall sketch how this result can be extended to positive sentences with arbitrary quantifier prefix.

Before we can describe the algorithm, we must introduce some notation. In the following, V denotes an infinite set of variables used by the first order languages under consideration. Let t be a $(\Sigma \cup \Delta)$-term. This term is called *pure* iff it is either a Σ-term or a Δ-term. An equation is pure iff it is an equation between pure terms of the same signature. A relational formula $p[s_1, \ldots, s_m]$ is pure iff $s_1, \ldots, s_m$ are pure terms of the signature of p. Now assume that t is a non-pure term whose topmost function symbol is in Σ. A subterm s of t is called *alien subterm* of t iff its topmost function symbol belongs to Δ and every proper superterm of s in t has its top symbol in Σ. Alien subterms of terms with top symbol in Δ are defined analogously. For a relational formula $p[s_1, \ldots, s_m]$, alien subterms are defined as follows: if s_i has a top symbol whose signature is different from the signature of p then s_i itself is an alien subterm; otherwise, any alien subterm of s_i is an alien subterm of $p[s_1, \ldots, s_m]$.

4.1 The Decomposition Algorithm

Let φ_0 be a positive existential $(\Sigma \cup \Delta)$-sentence. Without loss of generality, we may assume that φ_0 has the form $\exists \vec{u}_0 \; \gamma_0$, where γ_0 is a conjunction of atomic formulae. Indeed, since existential quantifiers distribute over disjunction, a sentence $\exists \vec{u}_0 \; (\gamma_1 \vee \gamma_2)$ is valid iff $\exists \vec{u}_0 \; \gamma_1$ or $\exists \vec{u}_0 \; \gamma_2$ is valid.

Step 1: Transform non-pure atomic formulae.

(1) Equations $s = t$ of γ_0 where s and t have topmost function symbols belonging to different signatures are replaced by (the conjunction of) two new equations $u = s, u = t$, where u is a new variable. The quantifier prefix is extended by adding an existential quantification for u.

(2) As a result, we may assign a unique label Σ or Δ to each atomic formula that is not an equation between variables. The label of an equation $s = t$ is the signature of the topmost function symbols of s and/or t. The label of a relational formula $p[s_1, \ldots, s_m]$ is the signature of p.

(3) Now alien subterms occurring in atomic formulae are successively replaced by new variables. For example, assume that $s = t$ is an equation in the current formula, and that s contains the alien subterm s_1. Let u be a variable not occurring in the current formula, and let s' be the term obtained from s by replacing s_1 by u. Then the original equation is replaced by (the conjunction of) the two equations $s' = t$ and $u = s_1$. The quantifier prefix is extended by adding an existential quantification for u. The equation $s' = t$ keeps the label of $s = t$, and the label of $u = s_1$ is the signature of the top

symbol of s_1. Relational atomic formulae with alien subterms are treated analogously. This process is iterated until all atomic formulae occurring in the conjunctive matrix are pure. It is easy to see that this is achieved after finitely many iterations.

Step 2: Remove atomic formulae without label.
Equations between variables occurring in the conjunctive matrix are removed as follows: If $u = v$ is such an equation then one removes $\exists u$ from the quantifier prefix and $u = v$ from the matrix. In addition, every occurrence of u in the remaining matrix is replaced by v. This step is iterated until the matrix contains no equations between variables.

Let φ_1 be the new sentence obtained this way. The matrix of φ_1 can be written as a conjunction $\gamma_{1,\Sigma} \wedge \gamma_{1,\Delta}$, where $\gamma_{1,\Sigma}$ is a conjunction of all atomic formulae from φ_1 with label Σ, and $\gamma_{1,\Delta}$ is a conjunction of all atomic formulae from φ_1 with label Δ. There are three different types of variables occurring in φ_1: shared variables occur both in $\gamma_{1,\Sigma}$ and in $\gamma_{1,\Delta}$; Σ-variables occur only in $\gamma_{1,\Sigma}$; and Δ-variables occur only in $\gamma_{1,\Delta}$. Let $\vec{u}_{1,\Sigma}$ be the tuple of all Σ-variables, $\vec{u}_{1,\Delta}$ be the tuple of all Δ-variables, and $\vec{u}_1$ be the tuple of all shared variables.[5] Obviously, φ_1 is equivalent to the sentence

$$\exists \vec{u}_1 \, (\exists \vec{u}_{1,\Sigma} \; \gamma_{1,\Sigma} \wedge \exists \vec{u}_{1,\Delta} \; \gamma_{1,\Delta}) \, .$$

The next two steps of the algorithm are nondeterministic, i.e., a given sentence is transformed into finitely many new sentences. Here the idea is that the original sentence is valid iff at least one of the new sentences is valid.

Step 3: Variable identification.
Choose (nondeterministically) a partition of the set of all shared variables. The variables in each class of the partition are "identified" with each other by choosing an element of the class as representative, and replacing in the sentence all occurrences of variables of the class by this representative. Quantifiers for replaced variables are removed.

Let $\exists \vec{u}_2 \, (\exists \vec{u}_{1,\Sigma} \; \gamma_{2,\Sigma} \wedge \exists \vec{u}_{1,\Delta} \; \gamma_{2,\Delta})$ denote one of the sentences obtained by Step 3.

Step 4: Choose signature labels and ordering.
We choose a label Σ or Δ for every (shared) variable in $\vec{u}_2$, and a linear ordering $<$ on these variables.

For each of the choices made in Step 3 and 4, the algorithm yields a pair (α, β) of sentences as output.

Step 5: Generate output sentences.
The sentence $\exists \vec{u}_2(\exists \vec{u}_{1,\Sigma} \; \gamma_{2,\Sigma} \wedge \exists \vec{u}_{1,\Delta} \; \gamma_{2,\Delta})$ is split into two sentences

$$\alpha = \forall \vec{v}_1 \exists \vec{w}_1 \ldots \forall \vec{v}_k \exists \vec{w}_k \exists \vec{u}_{1,\Sigma} \; \gamma_{2,\Sigma} \text{ and } \beta = \exists \vec{v}_1 \forall \vec{w}_1 \ldots \exists \vec{v}_k \forall \vec{w}_k \exists \vec{u}_{1,\Delta} \; \gamma_{2,\Delta}.$$

[5] The order in these tuples can be chosen arbitrarily.

Here $\vec{v}_1\vec{w}_1\ldots\vec{v}_k\vec{w}_k$ is the unique re-ordering of $\vec{u}_2$ along $<$. The variables $\vec{v}_i$ $(\vec{w}_i)$ are the variables with label Δ (label Σ).

Thus, the overall output of the algorithm is a finite set of pairs of sentences. Note that the sentences α and β are positive formulae, but they need no longer be existential positive formulae.

This algorithm is a straightforward adaptation of the decomposition algorithm described in [BS92] to existential positive formulae with equations and relational constraints. Note, however, that it optimizes the previous algorithm in one significant way: the nondeterministic steps—which are responsible for the NP-complexity of the algorithm—are applied only to shared variables and not to all variables occurring in the system. For the case of algorithms computing complete sets of unifiers, this optimization is already implicitly present in [Bou93]. Steps similar to Step 1, 3, and the labelling in Step 4 are present in most methods for combining unification algorithms. Nelson & Oppen's combination method for universal theories [NO79] explicitly uses Step 1, and implicitly, Step 3 is also present.

4.2 Correctness of the decomposition algorithm

First, we show soundness of the algorithm, i.e., if one of the output pairs is valid then the original sentence was valid.

Lemma 7. *$\mathcal{A} \odot \mathcal{B} \models \varphi_0$ if $\mathcal{A} \models \alpha$ and $\mathcal{B} \models \beta$ for some output pair (α, β).*

Proof. Since $\mathcal{A}^{\Sigma}$ and $\mathcal{A}^{\Sigma}_{\infty}$ are isomorphic Σ-structures, we know that $\mathcal{A}^{\Sigma}_{\infty} \models \alpha$. Accordingly, we also have $\mathcal{B}^{\Delta}_{\infty} \models \beta$. More precisely, this means

$$(*)\quad \mathcal{A}^{\Sigma}_{\infty} \models \forall\vec{v}_1\exists\vec{w}_1\ldots\forall\vec{v}_k\exists\vec{w}_k\exists\vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(\vec{v}_1,\vec{w}_1,\ldots,\vec{v}_k,\vec{w}_k,\vec{u}_{1,\Sigma})$$
$$(**)\quad \mathcal{B}^{\Delta}_{\infty} \models \exists\vec{v}_1\forall\vec{w}_1\ldots\exists\vec{v}_k\forall\vec{w}_k\exists\vec{u}_{1,\Delta}\ \gamma_{2,\Delta}(\vec{v}_1,\vec{w}_1,\ldots,\vec{v}_k,\vec{w}_k,\vec{u}_{1,\Delta}).$$

Because of the existential quantification over $\vec{v}_1$ in $(**)$, there exist elements $\vec{b}_1 \in \vec{B}_{\infty}$ such that

$$(***)\quad \mathcal{B}^{\Delta}_{\infty} \models \forall\vec{w}_1\ldots\exists\vec{v}_k\forall\vec{w}_k\exists\vec{u}_{1,\Delta}\ \gamma_{2,\Delta}(\vec{b}_1,\vec{w}_1,\ldots,\vec{v}_k,\vec{w}_k,\vec{u}_{1,\Delta}).$$

We consider $\vec{a}_1 := h_{\infty}(\vec{b}_1)$. Because of the universal quantification over $\vec{v}_1$ in $(*)$ we have

$$\mathcal{A}^{\Sigma}_{\infty} \models \exists\vec{w}_1\ldots\forall\vec{v}_k\exists\vec{w}_k\exists\vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(\vec{a}_1,\vec{w}_1,\ldots,\vec{v}_k,\vec{w}_k,\vec{u}_{1,\Sigma}).$$

Because of the existential quantification over $\vec{w}_1$ in this formula there exist elements $\vec{c}_1 \in \vec{A}_{\infty}$ such that

$$\mathcal{A}^{\Sigma}_{\infty} \models \forall\vec{v}_2\exists\vec{w}_2\ldots\forall\vec{v}_k\exists\vec{w}_k\exists\vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(\vec{a}_1,\vec{c}_1,\vec{v}_2,\vec{w}_2,\ldots,\vec{v}_k,\vec{w}_k,\vec{u}_{1,\Sigma}).$$

We consider $\vec{d}_1 := g_{\infty}(\vec{c}_1)$. Because of the universal quantification over $\vec{w}_1$ in $(***)$ we have

$$\mathcal{B}^{\Delta}_{\infty} \models \exists\vec{v}_2\forall\vec{w}_2\ldots\exists\vec{v}_k\forall\vec{w}_k\exists\vec{u}_{1,\Delta}\ \gamma_{2,\Delta}(\vec{b}_1,\vec{d}_1,\vec{v}_2,\vec{w}_2,\ldots,\vec{v}_k,\vec{w}_k,\vec{u}_{1,\Delta}).$$

Iterating this argument, we thus obtain

$$\mathcal{A}_\infty^\Sigma \models \exists \vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(\vec{a}_1, \vec{c}_1, \ldots, \vec{a}_k, \vec{c}_k, \vec{u}_{1,\Sigma}),$$
$$\mathcal{B}_\infty^\Delta \models \exists \vec{u}_{1,\Delta}\ \gamma_{2,\Delta}(\vec{b}_1, \vec{d}_1, \ldots, \vec{b}_k, \vec{d}_k, \vec{u}_{1,\Delta}),$$

where $\vec{a}_i = h_\infty(\vec{b}_i)$ and $\vec{d}_i = g_\infty(\vec{c}_i)$ (for $1 \leq i \leq k$). Since h_∞ is a $(\Sigma \cup \Delta)$-isomorphism that is the inverse of g_∞, we also know that

$$\mathcal{A}_\infty^\Delta \models \exists \vec{u}_{1,\Delta}\ \gamma_{2,\Delta}(\vec{a}_1, \vec{c}_1, \ldots, \vec{a}_k, \vec{c}_k, \vec{u}_{1,\Delta}).$$

It follows that

$$\mathcal{A}_\infty^{\Sigma\cup\Delta} \models \exists \vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(\vec{a}_1, \vec{c}_1, \ldots, \vec{a}_k, \vec{c}_k, \vec{u}_{1,\Sigma}) \wedge \exists \vec{u}_{1,\Delta}\ \gamma_{2,\Delta}(\vec{a}_1, \vec{c}_1, \ldots, \vec{a}_k, \vec{c}_k, \vec{u}_{1,\Delta}).$$

Obviously, this implies that $\mathcal{A} \odot \mathcal{B} \simeq \mathcal{A}_\infty^{\Sigma\cup\Delta} \models \exists \vec{u}_2\,(\exists \vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma} \wedge \exists \vec{u}_{1,\Delta}\ \gamma_{2,\Delta})$, i.e., one of the sentences obtained after Step 3 of the algorithm holds in $\mathcal{A} \odot \mathcal{B}$. It is easy to see that this implies that $\mathcal{A} \odot \mathcal{B} \models \varphi_0$. □

Next, we show completeness of the decomposition algorithm, i.e., if the input sentence was valid then there exists a valid output pair.

Lemma 8. *If $\mathcal{A} \odot \mathcal{B} \models \varphi_0$ then $\mathcal{A} \models \alpha$ and $\mathcal{B} \models \beta$ for some output pair (α, β).*

Proof. Assume that $\mathcal{A} \odot \mathcal{B} \simeq \mathcal{B}_\infty^{\Sigma\cup\Delta} \models \exists \vec{u}_0 \gamma_0$. Obviously, this implies that $\mathcal{B}_\infty^{\Sigma\cup\Delta} \models \exists \vec{u}_1\,(\exists \vec{u}_{1,\Sigma}\ \gamma_{1,\Sigma}(\vec{u}_1, \vec{u}_{1,\Sigma}) \wedge \exists \vec{u}_{1,\Delta}\ \gamma_{1,\Delta}(\vec{u}_1, \vec{u}_{1,\Delta}))$, i.e., $\mathcal{B}_\infty^{\Sigma\cup\Delta}$ satisfies the sentence that is obtained after Step 2 of the decomposition algorithm. Thus there exists an assignment $\nu : V \to B_\infty$ such that $\mathcal{B}_\infty^{\Sigma\cup\Delta} \models \exists \vec{u}_{1,\Sigma}\ \gamma_{1,\Sigma}(\nu(\vec{u}_1), \vec{u}_{1,\Sigma}) \wedge \exists \vec{u}_{1,\Delta}\ \gamma_{1,\Delta}(\nu(\vec{u}_1), \vec{u}_{1,\Delta})$.

In Step 3 of the decomposition algorithm we identify two shared variables u and u' of $\vec{u}_1$ if, and only if, $\nu(u) = \nu(u')$. With this choice,

$$\mathcal{B}_\infty^{\Sigma\cup\Delta} \models \exists \vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(\nu(\vec{u}_2), \vec{u}_{1,\Sigma}) \wedge \exists \vec{u}_{1,\Delta}\ \gamma_{2,\Delta}(\nu(\vec{u}_2), \vec{u}_{1,\Delta}),$$

and all components of $\nu(\vec{u}_2)$ are distinct.

In Step 4, a shared variable u in $\vec{u}_2$ is labeled with Δ if $\nu(u) \in B_\infty \setminus (\bigcup_{i=1}^\infty Y_i)$, and with Σ otherwise. In order to choose the linear ordering on the shared variables, we partition the range B_∞ of ν as follows:

$$B_0,\ \ Y_1,\ \ B_1 \setminus (B_0 \cup Y_1),\ \ Y_2,\ \ B_2 \setminus (B_1 \cup Y_2),\ \ Y_3,\ \ B_3 \setminus (B_2 \cup Y_3),\ \ldots$$

Now, let $\vec{v}_1, \vec{w}_1, \ldots, \vec{v}_k, \vec{w}_k$ be a re-ordering of the tuple $\vec{u}_2$ such that the following holds:

1. The tuple $\vec{v}_1$ contains exactly the shared variables whose ν-images are in B_0.
2. For all $i, 1 \leq i \leq k$, the tuple $\vec{w}_i$ contains exactly the shared variables whose ν-images are in Y_i.
3. For all $i, 1 < i \leq k$, the tuple $\vec{v}_i$ contains exactly the shared variables whose ν-images are in $B_{i-1} \setminus (B_{i-2} \cup Y_{i-1})$.

Obviously, this implies that the variables in the tuples $\vec{w}_i$ have label Σ, whereas the variables in the tuples $\vec{v}_i$ have label Δ. Note that some of these tuples may be of dimension 0. The re-ordering determines the linear ordering we choose in Step 4. Let

$$\alpha = \forall\vec{v}_1\exists\vec{w}_1\ldots\forall\vec{v}_k\exists\vec{w}_k\exists\vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma} \quad \text{and} \quad \beta = \exists\vec{v}_1\forall\vec{w}_1\ldots\exists\vec{v}_k\forall\vec{w}_k\exists\vec{u}_{1,\Delta}\ \gamma_{2,\Delta}$$

be the output pair that is obtained by these choices. Let $\vec{y}_i := \nu(\vec{w}_i) \in \vec{Y}$ and $\vec{b}_i := \nu(\vec{v}_i) \in \vec{B}_\infty$. The sequence $\vec{b}_1, \vec{y}_1, \ldots, \vec{b}_k, \vec{y}_k$ satisfies Condition 2 of Lemma 5 for $\varphi = \exists\vec{u}_{1,\Delta}\ \gamma_{2,\Delta}$, the structure $\mathcal{B}_\infty^\Delta$, and appropriate sets $Z_1, \ldots, Z_k$ (see [BS94]). Thus, we obtain $\mathcal{B} \simeq \mathcal{B}_\infty^\Delta \models \beta$. In order to show $\mathcal{A} \models \alpha$, we use the fact that $h_\infty : \mathcal{B}_\infty \to \mathcal{A}_\infty$ is a $(\Sigma \cup \Delta)$-isomorphism. Thus, $\mathcal{B}_\infty^{\Sigma\cup\Delta} \models \exists\vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(\nu(\vec{u}_2), \vec{u}_{1,\Sigma})$ implies that $\mathcal{A}_\infty^\Sigma \models \exists\vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}(h_\infty(\nu(\vec{u}_2)), \vec{u}_{1,\Sigma})$.

Let $\vec{x}_i := h_\infty(\vec{b}_i) = h_\infty(\nu(\vec{v}_i))$ and $\vec{a}_i := h_\infty(\vec{y}_i) = h_\infty(\nu(\vec{w}_i))$ (for $i = 1, \ldots, k$). In [BS94] it is shown that the sequence $\vec{x}_1, \vec{a}_1, \ldots, \vec{x}_k, \vec{a}_k$ satisfies Condition 2 of Lemma 5 for $\varphi = \exists\vec{u}_{1,\Sigma}\ \gamma_{2,\Sigma}$, the structure $\mathcal{A}_\infty^\Sigma$, and appropriate sets $Z'_1, \ldots, Z'_k$. Thus, $\mathcal{A} \simeq \mathcal{A}_\infty^\Sigma \models \alpha$. □

The two lemmas obviously imply the next theorem.

Theorem 9. *Let $\mathcal{V}(E)$ be a Σ-variety and $\mathcal{V}(F)$ be a Δ-variety for disjoint signatures Σ and Δ. The positive existential theory of the $(\Sigma\cup\Delta)$-variety $\mathcal{V}(E\cup F)$ is decidable, provided that the positive theories of $\mathcal{V}(E)$ and of $\mathcal{V}(F)$ are decidable.*

If the signatures contain no predicate symbols, this theorem is a reformulation of Theorem 2.1 of [BS92]. What is new here is the algebraic proof method and the fact that relational constraints can be treated as well.

4.3 Decision Procedures for Positive Theories

A disadvantage of Theorem 9 is that it does not show modularity of decidability of the positive theory of varieties of structures. Indeed, the prerequisites of the theorem (decidability of the *full* positive theories of $\mathcal{V}(E)$ and $\mathcal{V}(F)$) are stronger than its consequence (decidability of the *existential* positive theory of $\mathcal{V}(E\cup F)$).

In [BS94] we describe an algorithm that can be used to reduce decidability of the *full* positive theory of $\mathcal{V}(E \cup F)$ to decision procedures for the positive theories of $\mathcal{V}(E)$ and $\mathcal{V}(F)$. The main idea is to transform positive sentences (with arbitrary quantifier prefix) into existential positive sentences by Skolemizing the universally quantified variables.[6] In addition to the theories E and F one thus obtains a free theory (for the new Skolem functions). In principle, the decomposition algorithm for existential positive sentences is now applied twice to decompose the input sentence into three positive sentences α, β, ρ, whose validity must respectively be decided in E, F, and the free theory. Note that it is well-known that the whole first-order theory of absolutely free term algebras is decidable [Mal71, Mah88, CL89].

[6] We are Skolemizing *universally* quantified variables since we are interested in validity of the sentence and not in satisfiability.

Correctness of this way of proceeding can be shown with the help of the following lemma, which exhibits an interesting connection between Skolemization and amalgamation with an absolutely free algebra (see [BS94] for the proof).

Lemma 10. *Let $\mathcal{A}$ be a Σ-structure that is free in $\mathcal{V}(E)$ over the countably infinite set of generators X, and let γ be a positive Σ-sentence. Suppose that the (positive) existential sentence γ' is obtained from γ via Skolemization of the universally quantified variables in γ, introducing the set of Skolem function symbols Γ. Then $\mathcal{A} \models \gamma$ if, and only if, $\mathcal{A} \odot \mathcal{T}(\Gamma, X) \models \gamma'$.*

Thus, we obtain the desired modularity result:

Theorem 11. *Let $\mathcal{V}(E)$ be a Σ-variety and $\mathcal{V}(F)$ be a Δ-variety for disjoint signatures Σ and Δ. The positive theory of the $(\Sigma \cup \Delta)$-variety $\mathcal{V}(E \cup F)$ is decidable, provided that the positive theories of $\mathcal{V}(E)$ and of $\mathcal{V}(F)$ are decidable.*

5 Conclusion and Outlook

We have presented an abstract algebraic approach to the problem of combining constraint solvers for constraint languages over disjoint signatures. The constraints that can be handled this way are built from atomic equational *and relational* constraints with the help of conjunction, disjunction, and both universal and existential quantifiers. Solvability means validity of such (closed) constraint formulae in a free structure, or equivalently in a variety of structures.

Simple examples of free structures with a non-trivial relational part are (absolutely free) term algebras that are equipped with an ordering that is invariant under substitution, such as the lexicographic path ordering or the subterm ordering. For our combination result to apply, however, the positive theory of these structures must be decidable. For a total lexicographic path ordering, this is not the case. For the subterm ordering, the existential theory is decidable, but the full first-order theory is undecidable [CT94]. Decidability of the positive theory is still an open problem. For partial lexicographic path orderings, even decidability of the existential theory is unknown.

Combination of constraint solving techniques in the presence of predicate symbols other than equality have independently been considered by H. Kirchner and Ch. Ringeissen [KR94]. However, their approach is based on the rewriting and abstraction techniques mentioned in the introduction (see, e.g., [BS92, Bou93]). Consequently, the interpretation of the predicate symbols in the combined structure is defined in a rather technical way, and it is not a priori clear what this definition means in an intuitive algebraic sense. We conjecture that, for free structures, the combined structure of [KR94] coincides with our free amalgamated product.

We are currently working on a generalization of the notion of "combinable structure" that considerably extends the notion of a "free structure." An example of a structure that is not a free structure, but nevertheless satisfies the generalized combinability condition, is the algebra of rational trees.

References

[Bou93] A. Boudet. Combining unification algorithms. *J. Symbolic Computation*, 16:597–626, 1993.

[BS92] F. Baader and K.U. Schulz. Unification in the union of disjoint equational theories: Combining decision procedures. In *Proceedings of CADE-11*, LNCS 607, 1992.

[BS93] F. Baader and K.U. Schulz. Unification in the union of disjoint equational theories: Combining decision procedures, 1993. Extended version, submitted for publication.

[BS93a] F. Baader and K.U. Schulz. Combination techniques and decision problems for disunification. In *Proceedings of RTA-93*, LNCS 690, 1993.

[BS94] F. Baader and K.U. Schulz. *Combination of Constraint Solving Techniques: An Algebraic Point of View.* Research Report CIS-Rep-94-75, CIS, University Munich, 1994. This report is available via anonymous ftp from "cantor.informatik.rwth-aachen.de" in the directory "pub/papers."

[Bür91] H.-J. Bürckert. *A Resolution Principle for a Logic with Restricted Quantifiers*, LNCS 568, 1991.

[Coh65] P.M. Cohn. *Universal Algebra.* Harper & Row, New York, 1965.

[CL89] H. Comon and P. Lescanne. Equational problems and disunification. *J. Symbolic Computation*, 7:371–425, 1989.

[CT94] H. Comon and R. Treinen. Ordering constraints on trees. In *Colloquium on Trees in Algebra and Programming (CAAP)*, LNCS, 1994.

[Der87] N. Dershowitz. Termination of rewriting. *J. Symbolic Computation*, 3:69–116, 1987.

[DKR94] E. Domenjoud, F. Klay, and Ch. Ringeissen. Combination techniques for non-disjoint theories. In *Proceedings of CADE-12*, LNCS 814, 1994.

[KK89] C. Kirchner and H. Kirchner. Constrained equational reasoning. In *Proceedings of SIGSAM 1989 International Symposium on Symbolic and Algebraic Computation.* ACM Press, 1989.

[KR94] H. Kirchner and Ch. Ringeissen. Combining symbolic constraint solvers on algebraic domains. *J. Symbolic Computation*, 18(2):113–155, 1994.

[Mah88] M.J. Maher. Complete axiomatizations of the algebras of finite, rational and infinite trees. In *Proceedings of LICS'88*, IEEE Computer Society, 1988.

[Mal71] A.I. Mal'cev. *The Metamathematics of Algebraic Systems*, volume 66 of *Studies in Logic and the Foundation of Mathematics.* North Holland, Amsterdam, London, 1971.

[Mal73] A.I. Mal'cev. *Algebraic Systems*, volume 192 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen.* Springer, Berlin, 1973.

[NO79] G. Nelson and D.C. Oppen. Simplification by cooperating decision procedures. *ACM TOPLAS*, 1(2):245–257, 1979.

[Rin92] Ch. Ringeissen. Unification in a combination of equational theories with shared constants and its application to primal algebras. In *Proceedings of LPAR'92*, LNCS 624, 1992.

[NiR94] R. Nieuwenhuis and A. Rubio, "AC-superposition with constraints: No AC-unifiers needed," in: *Proceedings CADE-12*, Springer LNAI 814, 1994.

[SS89] M. Schmidt-Schauß. Unification in a combination of arbitrary disjoint equational theories. *J. Symbolic Computation*, 8(1,2):51–99, 1989.

[Wea93] N. Weaver. Generalized varieties. *Algebra Universalis*, 30:27–52, 1993.