A

# Quantifier-Free Interpolation in Combinations of Equality Interpolating Theories

ROBERTO BRUTTOMESSO, Università degli Studi - Milano
SILVIO GHILARDI, Università degli Studi - Milano
SILVIO RANISE, FBK - Trento

The use of interpolants in verification is gaining more and more importance. Since theories used in applications are usually obtained as (disjoint) combinations of simpler theories, it is important to modularly re-use interpolation algorithms for the component theories. We show that a sufficient and necessary condition to do this for quantifier-free interpolation is that the component theories have the 'strong (sub-)amalgamation' property. Then, we provide an equivalent syntactic characterization and show that such characterization covers most theories commonly employed in verification. Finally, we design a combined quantifier-free interpolation algorithm capable of handling both convex and non-convex theories; this algorithm subsumes and extends most existing work on combined interpolation.

Categories and Subject Descriptors: F.4.1 [**Mathematical Logic**]: Computational Logic

General Terms: Algorithms, Theory, Verification

Additional Key Words and Phrases: Combined Interpolation, Craig Interpolation Theorem, Satisfiability Modulo Theories, Strong Amalgamability

**ACM Reference Format:**
Roberto Bruttomesso, Silvio Ghilardi, and Silvio Ranise, 2013. Quantifier-Free Interpolation in Combinations of Equality Interpolating Theories *ACM Trans. Comput. Logic* V, N, Article A (January YYYY), 34 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

## 1. INTRODUCTION

Craig's interpolation theorem [Craig 1957] applies to first order logic formulæ and states that whenever the sequent $\phi \wedge \psi \Rightarrow \bot$ is valid,[1] then it is possible to derive a formula $\theta$ such that

$(i)$. $\phi \Rightarrow \theta$ is valid,
$(ii)$. $\theta \wedge \psi \Rightarrow \bot$ is valid, and

[1]Strictly speaking, Craig's original formulation talks about valid implications, not about unsatisfiable conjunctions; this does not make a substantial difference in the context of this paper, but see [Kovcs and Voronkov 2009] for a thorough discussion.

The second author would like to acknowledge the support of the PRIN 2010-2011 project "Logical Methods for Information Management" funded by the Italian Ministry of Education, University and Research (MIUR). The work of the third author is partially supported by the "Automated Security Analysis of Identity and Access Management Systems (SIAM)" project funded by Provincia Autonoma di Trento in the context of the "team 2009 - Incoming" COFUND action of the European Commission (FP7).
Author's addresses: R. Bruttomesso, Dipartimento di Scienze dell'Informazione, via Comelico 39/41, 20135 Milano, Italy; S. Ghilardi, Dipartimento di Matematica, via C. Saldini 50, 20123 Milano, Italy; S. Ranise, FBK (Fondazione Bruno Kessler), Povo - Via Sommarive 18, 38123 Trento, Italy.

DOI: http://dx.doi.org/10.1145/0000000.0000000

$(iii)$. $\theta$ is defined over the common symbols of $\phi$ and $\psi$.

After the seminal work of McMillan (see, e.g., [McMillan 2005b]), Craig's interpolation has become an important technique in verification. Intuitively, the interpolant $\theta$ can be seen as an over-approximation of $\phi$ with respect to $\psi$. Using this observation, algorithms for computing interpolants are more and more used in verification. For example, in the abstraction-refinement phase of software model checking [Henzinger and R. Jhala 2004], interpolants are used to compute increasingly precise over-approximations of the set of reachable states. Of particular importance for verification techniques are those algorithms capable of computing *quantifier-free* interpolants in presence of some background theory. This is so because several symbolic verification problems are formalized by representing sets of states and transitions as quantifier-free formulae. Quantifier-free interpolants might be used also when verification tools need to handle quantified formulae [McMillan 2011; Alberti et al. 2012a; 2012b].

Unfortunately, Craig's interpolation theorem does not guarantee that it is always possible to compute quantifier-free interpolants when reasoning *modulo a first-order theory*: in fact, for certain first-order theories, it is known that quantifiers must occur in interpolants of quantifier-free formulae [Kapur et al. 2006]. Even when quantifier-free interpolants exist for single theories, this might not be the case when considering their combinations: a typical example of this phenomenon is Presburger arithmetic that loses quantifier-free interpolants when combined with the pure theory of equality with uninterpreted symbols [Brillout et al. 2010]. Since verification techniques frequently require to reason in combinations of theories, methods to modularly combine available interpolation algorithms are indeed desirable.

This paper studies the *modularity of quantifier-free interpolation*. The starting point is the well-known fact [Bacsich 1975] that quantifier-free interpolation (for universal theories) is equivalent to the model-theoretic property of *amalgamability*. Intuitively, a theory has the amalgamation property if any two structures $\mathcal{M}_1, \mathcal{M}_2$ in its class of models, sharing a common sub-model $\mathcal{M}_0$, can be regarded as sub-structures of a larger model $\mathcal{M}$, called the amalgamated model. This property is not sufficient to derive a modularity result for quantifier-free interpolation. As shown below, a stronger notion is needed, called *strong amalgamability* [Jónsson 1956]; this notion has been thoroughly analyzed in universal algebra and category theory [Ringel 1972; Kiss et al. 1982]. A theory has the strong amalgamation property if elements from the supports of $\mathcal{M}_1, \mathcal{M}_2$ not belonging to the support of $\mathcal{M}_0$ *cannot be identified* in the amalgamated model $\mathcal{M}$. An example of an amalgamable but not strongly amalgamable theory is the theory of fields: let $\mathcal{M}_0$ be a real field and $\mathcal{M}_1, \mathcal{M}_2$ be two copies of the complex numbers, the imaginary unit in $\mathcal{M}_1$ must be identified with the imaginary unit of $\mathcal{M}_2$ (or with its opposite) in any amalgamating field $\mathcal{M}$ since the polynomial $x^2 + 1$ cannot have more than two roots. More examples will be discussed below and many more can be found in [Kiss et al. 1982].

The first contribution of this paper is to show that *strong amalgamability is precisely what is needed to guarantee the modularity of quantifier-free interpolation*. Formally, we prove the following two properties:[2]

*(*a*)*. if $T_1$ and $T_2$ are signature disjoint, both stably infinite and strongly amalgamable, then $T_1 \cup T_2$ is also strongly amalgamable and hence quantifier-free interpolating and

[2]Here, for simplicity, we assume theories to be universal. In the rest of the paper, we will generalize our results to arbitrary theories: the generalization simply requires to replace (strong) amalgamability with (strong) *sub*-amalgamability.

*(*b*).* a theory $T$ is strongly amalgamable iff the disjoint union of $T$ with the theory $\mathcal{EUF}$ of equality with uninterpreted symbols has quantifier-free interpolation (Section 3).

Notice how the first two requirements in property (*a*) guarantee also the correctness of the Nelson-Oppen method [Nelson and Oppen 1979], whose importance for combined satisfiability problems is well-known.

Our proof of property (*a*) is non-constructive. As a consequence, it does not provide the basis to design an algorithm computing quantifier-free interpolants in combinations of theories. To overcome this problem, the second contribution of the paper consists of identifying the class of *equality interpolating theories*. A theory $T$ is equality interpolating when it is possible to compute terms that are equal to the variables occurring in disjunctions of equalities entailed (modulo $T$) by pairs of quantifier-free formulae. Most importantly, we show that *equality interpolation is equivalent to strong amalgamation* (Section 4). In other words, equality interpolation is a more operational notion of strong amalgamability, that can be exploited to design an algorithm to compute quantifier-free interpolants in combination of theories. Before giving the algorithm, we show that several theories—that are well-known in the literature and considered to be useful in verification—are equality interpolating; e.g., the theory of recursively defined data structures [Oppen 1980], Integer Difference Logic, Unit-Two-Variable-Per-Inequality, and Integer Linear Arithmetic with division-by-$n$ [Brillout et al. 2011]. This result is shown by proving that universal theories that also admit elimination of quantifiers [Enderton 1972] are equality interpolating (Section 4.1). The third contribution of the paper is *a combination algorithm for the computation of quantifier-free interpolants from finite sets of quantifier-free formulae in unions of signature disjoint, stably infinite, and equality interpolating theories* (Section 6). The algorithm uses as sub-modules the interpolation algorithms of the component theories and is based on a sequence of syntactic manipulations organized in groups of transformations modelled after a non-deterministic version of the Nelson-Oppen combination schema (see, e.g., [Tinelli and Harandi 1996]).

The last contribution of the paper is to *clarify some issues arising from recent and less recent works in the literature about combination of interpolation algorithms*.

First, we relate our notion of equality interpolation with that in [Yorsh and Musuvathi 2005; 2004]. We prove that the two notions coincide for convex theories, but this is no more the case in the non-convex case (Section 5). This implies that our notion of equality interpolating theory is weaker and thus our combination results are *stronger* than those in [Yorsh and Musuvathi 2005; 2004].

Second, we give a (formal) counterexample showing that—as expected—the 'convex' equality interpolating condition in [Yorsh and Musuvathi 2005] is not sufficient to guarantee combined quantifier-free interpolation in non convex theories (Section 5.3).

Third, we prove that the general formulation of the interpolation property used in formal verification can be reduced to combined interpolation in the quantifier-free case (Section 3.3).

Finally, we show the close connection of our equality interpolation condition with *Beth definability property*, a standard notion in model theory that has also interesting reformulation in universal algebra (Section 5.4). Via Beth definability, we are able to compare our results with characterizations of strong amalgamability from the algebraic literature, such as those in, e.g., [Kiss et al. 1982].

All these observations seem to suggest that the notion of equality interpolation given in the paper is the "right one:"

— it is more general than that in [Yorsh and Musuvathi 2005; 2004],

— it is both sufficient and *necessary* (in the sense of properties (a) and (b) stated above) to guarantee the modularity of quantifier-free interpolation, and
— it is closely related to the standard model theoretic notion of Beth definability so that it is possible to perform a precise comparison with other characterizations of strong amalgamability.

*Plan of the paper.* Section 2 fixes formal notation and gives basic notions. Section 3 states and proves results relating combined quantified free interpolation and strong amalgamation. Section 4 gives syntactic characterizations and supplies examples via quantifier elimination of universal axioms. Section 5 discusses the relationships with previous works on interpolation in combinations of theories. Section 5.4 covers the related model-theoretic notion of Beth definability. Section 6 describes our interpolation algorithm for combination of theories. Finally, Section 7 concludes by discussing our results in the context of previous works in the field of interpolation algorithms.

## 2. FORMAL PRELIMINARIES

We assume the usual syntactic (e.g., signature, variable, term, atom, literal, formula, and sentence) and semantic (e.g., structure, sub-structure, truth, satisfiability, and validity) notions of first-order logic (see, e.g., [Enderton 1972]). The equality symbol "$=$" is included in all signatures considered below. For clarity, we shall use "$\equiv$" in the meta-theory to express the syntactic identity between two symbols or two strings of symbols. Notations like $E(\underline{x})$ mean that the expression (term, literal, formula, etc.) $E$ contains free variables only from the tuple $\underline{x}$. A 'tuple of variables' is a list of variables without repetitions and a 'tuple of terms' is a list of terms (possibly with repetitions). Finally, whenever we use a notation like $E(\underline{x}, \underline{y})$ we implicitly assume not only that both the $\underline{x}$ and the $\underline{y}$ are pairwise distinct, but also that $\underline{x}$ and $\underline{y}$ are disjoint. A formula is *universal* (*existential*) iff it is obtained from a quantifier-free formula by prefixing it with a string of universal (existential, resp.) quantifiers.

*Theories and satisfiability modulo theory.* A *theory* $T$ is a pair $(\Sigma, Ax_T)$, where $\Sigma$ is a signature and $Ax_T$ is a set of $\Sigma$-sentences, called the *axioms* of $T$ (we shall sometimes write directly $T$ for $Ax_T$). The *models* of $T$ are those $\Sigma$-structures in which all the sentences in $Ax_T$ are true. A $\Sigma$-formula $\phi$ is *$T$-satisfiable* (or $T$-consistent) if there exists a model $\mathcal{M}$ of $T$ such that $\phi$ is true in $\mathcal{M}$ under a suitable assignment $\mathtt{a}$ to the free variables of $\phi$ (in symbols, $(\mathcal{M}, \mathtt{a}) \models \phi$); it is *$T$-valid* (in symbols, $T \vdash \varphi$) if its negation is $T$-unsatisfiable or, equivalently, $\varphi$ is provable from the axioms of $T$ in a complete calculus for first-order logic. A theory $T = (\Sigma, Ax_T)$ is *universal* iff there is a theory $T' = (\Sigma, Ax_{T'})$ such that all sentences in $Ax_{T'}$ are universal and the sets of $T$-valid and $T'$-valid sentences coincide. A formula $\varphi_1$ *$T$-entails* a formula $\varphi_2$ if $\varphi_1 \rightarrow \varphi_2$ is *$T$-valid* (in symbols, $\varphi_1 \vdash_T \varphi_2$ or simply $\varphi_1 \vdash \varphi_2$ when $T$ is clear from the context). If $\Gamma$ is a set of formulæ and $\phi$ a formula, $\Gamma \vdash_T \phi$ means that there are $\gamma_1, \ldots, \gamma_n \in \Gamma$ such that $\gamma_1 \wedge \cdots \wedge \gamma_n \vdash_T \phi$. The *satisfiability modulo the theory* $T$ (SMT($T$)) *problem* amounts to establishing the $T$-satisfiability of quantifier-free $\Sigma$-formulæ. Some theories have special names, these names are becoming standard in SMT-literature; we shall recall them during the paper, here we just fix the notation for the pure equality theory. We shall call $\mathcal{EUF}(\Sigma)$ the pure equality theory in the signature $\Sigma$; we may also use just $\mathcal{EUF}$ in case there is no need to specify the signature $\Sigma$ (notice that $\Sigma$ is arbitrary and, as such, it may contain only function symbols, only predicate symbols or both predicate and function symbols).

A theory $T$ admits *quantifier-elimination* iff for every formula $\phi(\underline{x})$ there is a quantifier-free formula $\phi'(\underline{x})$ such that $T \vdash \phi \leftrightarrow \phi'$.

*Embeddings, sub-structures, and diagrams.* The support (i.e. the underlying set) of a structure $\mathcal{M}$ is denoted with $|\mathcal{M}|$. An embedding is a homomorphism that preserves and reflects relations and operations (see, e.g., [Chang and Keisler 1990]). Formally, a $\Sigma$-*embedding* (or, simply, an embedding) between two $\Sigma$-structures $\mathcal{M}$ and $\mathcal{N}$ is any mapping $\mu : |\mathcal{M}| \longrightarrow |\mathcal{N}|$ satisfying the following three conditions: (a) it is a injective function; (b) it is an algebraic homomorphism, that is for every $n$-ary function symbol $f$ and for every $a_1, \ldots, a_n \in |\mathcal{M}|$, we have $f^{\mathcal{N}}(\mu(a_1), \ldots, \mu(a_n)) = \mu(f^{\mathcal{M}}(a_1, \ldots, a_n))$; (c) it preserves and reflects interpreted predicates, i.e. for every $n$-ary predicate symbol $P$, we have $(a_1, \ldots, a_n) \in P^{\mathcal{M}}$ iff $(\mu(a_1), \ldots, \mu(a_n)) \in P^{\mathcal{N}}$. If $|\mathcal{M}| \subseteq |\mathcal{N}|$ and the embedding $\mu : \mathcal{M} \longrightarrow \mathcal{N}$ is just the identity inclusion $|\mathcal{M}| \subseteq |\mathcal{N}|$, we say that $\mathcal{M}$ is a *substructure* of $\mathcal{N}$ or that $\mathcal{N}$ is a *superstructure* of $\mathcal{M}$. As it is well-known, the truth of a universal (resp. existential) sentence is preserved through substructures (resp. superstructures). The substructure *generated by* a subset $X$ of the support of a structure $\mathcal{M}$ is the smallest substructure of $\mathcal{M}$ containing $X$ in its support.

Given a signature $\Sigma$ and a $\Sigma$-structure $\mathcal{A}$, we indicate with $\Delta_\Sigma(\mathcal{A})$ the *diagram* of $\mathcal{A}$: this is the set of sentences obtained by first expanding $\Sigma$ with a fresh constant $\bar{a}$ for every element $a$ from $|\mathcal{A}|$ and then taking the set of ground $\Sigma \cup |\mathcal{A}|$-literals which are true in $\mathcal{A}$ (under the natural expanded interpretation mapping $\bar{a}$ to $a$).[3] An easy but nevertheless important basic result (to be frequently used in our proofs), called *Robinson Diagram Lemma* [Chang and Keisler 1990], says that, given any $\Sigma$-structure $\mathcal{B}$, there is an embedding $\mu : \mathcal{A} \longrightarrow \mathcal{B}$ iff $\mathcal{B}$ can be expanded to a $\Sigma \cup |\mathcal{A}|$-structure in such a way that it becomes a model of $\Delta_\Sigma(\mathcal{A})$.

*Combinations of theories.* A theory $T$ is *stably infinite* iff every $T$-satisfiable quantifier-free formula (from the signature of $T$) is satisfiable in an infinite model of $T$. By compactness, it is possible to show that $T$ is stably infinite iff every model of $T$ embeds into an infinite one (see for instance [Ghilardi 2004]). A theory $T$ is *convex* iff for every conjunction of literals $\delta$, if $\delta \vdash_T \bigvee_{i=1}^{n} x_i = y_i$ then $\delta \vdash_T x_i = y_i$ holds for some $i \in \{1, ..., n\}$.

Let $T_i$ be a stably-infinite theory over the signature $\Sigma_i$ such that the $SMT(T_i)$ problem is decidable for $i = 1, 2$ and $\Sigma_1$ and $\Sigma_2$ are disjoint (i.e. the only shared symbol is equality). Under these assumptions, the Nelson-Oppen combination method [Nelson and Oppen 1979] tells us that the SMT problem for the combination $T_1 \cup T_2$ of the theories $T_1$ and $T_2$ (i.e. the union of their axioms) is decidable.

*Interpolation properties.* Craig's interpolation theorem [Chang and Keisler 1990] roughly states that if a formula $\phi$ implies a formula $\psi$ then there is a third formula $\theta$, called an interpolant, such that $\phi$ implies $\theta$, $\theta$ implies $\psi$, and every non-logical symbol in $\theta$ occurs both in $\phi$ and $\psi$. In this paper, we are interested to specialize this result to the computation of quantifier-free interpolants modulo (combinations of) theories.

*Definition* 2.1. [Plain quantifier-free interpolation] A theory $T$ *admits (plain) quantifier-free interpolation* (or, equivalently, *has quantifier-free interpolants*) iff for every pair of quantifier-free formulae $\phi, \psi$ such that $\psi \wedge \phi$ is $T$-unsatisfiable, there exists a quantifier-free formula $\theta$, called an *interpolant*, such that: (i) $\psi$ $T$-entails $\theta$, (ii) $\theta \wedge \phi$ is $T$-unsatisfiable, and (iii) only the variables occurring in both $\psi$ and $\phi$ occur in $\theta$.

In verification, the following extension of the above definition is considered more useful.

[3] As usual in model theory books, we won't distinguish anymore an element $a \in |\mathcal{A}|$ from its name $\bar{a}$ in the expanded language $\Sigma \cup |\mathcal{A}|$.

*Definition* 2.2. [General quantifier-free interpolation] Let $T$ be a theory in a signature $\Sigma$; we say that $T$ has the *general quantifier-free interpolation property* iff for every signature $\Sigma'$ (disjoint from $\Sigma$) and for every pair of ground $\Sigma \cup \Sigma'$-formulæ $\phi, \psi$ such that $\phi \wedge \psi$ is $T$-unsatisfiable,[4] there is a ground formula $\theta$ such that: (i) $\phi$ $T$-entails $\theta$; (ii) $\theta \wedge \psi$ is $T$-unsatisfiable; (iv) all predicate, constants and function symbols from $\Sigma'$ occurring in $\theta$ also occur in $\phi$ and $\psi$.

By replacing free variables with free constants, it should be clear that the general quantifier-free interpolation property (Definition 2.2) implies the plain quantifier-free interpolation property (Definition 2.1). We shall first investigate plain quantifier-free interpolation and then (Section 3.3) relate general quantifier-free interpolation with interpolation in combinations of theories (in particular, with interpolation in combinations with $\mathcal{EUF}$).

## 3. STRONG AMALGAMATION AND QUANTIFIER-FREE INTERPOLATION

Amalgamability and strong amalgamability are usually formulated for universal theories (see, e.g., [Kiss et al. 1982] for a survey). In order to define suitable extensions to arbitrary first order theories, it is important to observe that a substructure of a model of a non-universal theory need not be a model of the theory. Thus, the extensions, that we would like to introduce, must take into account substructures which are not necessarily submodels. We call such extensions 'sub-amalgamability' and 'strong sub-amalgamability,' respectively, and formally define them as follows.

*Definition* 3.1. A theory $T$ has the *sub-amalgamation property* iff whenever we are given models $\mathcal{M}_1$ and $\mathcal{M}_2$ of $T$ and a common substructure $\mathcal{A}$ of them, there exists a further model $\mathcal{M}$ of $T$ endowed with embeddings $\mu_1 : \mathcal{M}_1 \longrightarrow \mathcal{M}$ and $\mu_2 : \mathcal{M}_2 \longrightarrow \mathcal{M}$ whose restrictions to $|\mathcal{A}|$ coincide.[5]

A theory $T$ has the *strong sub-amalgamation property* if the above embeddings $\mu_1, \mu_2$ and the above model $\mathcal{M}$ can be chosen so to satisfy the following additional condition: if for some $m_1, m_2$ we have $\mu_1(m_1) = \mu_2(m_2)$, then there exists an element $a$ in $|\mathcal{A}|$ such that $m_1 = a = m_2$.

### 3.1. Amalgamation and interpolation

The result of [Bacsich 1975] relating amalgamation and quantifier-free interpolation in universal theories can be easily extended to arbitrary theories by replacing amalgamation with sub-amalgamation. This extension is stated in Theorem 3.3 below. For the sake of completeness, we give a proof. Preliminarily, we need the following lemma.

LEMMA 3.2. *Let $T$ be a theory in a signature $\Sigma$ and $\underline{a}, \underline{b}, \underline{c}$ tuples of (distinct) free constants. Furthermore, let $\Theta_1, \Theta_2$ be sets of ground formulae such that*

— *at most the free constants $\underline{a}, \underline{c}$ occur in $\Theta_1$;*
— *at most the free constants $\underline{b}, \underline{c}$ occur in $\Theta_2$;*
— *there exists no ground formula $\theta(\underline{c})$ with $\Theta_1 \vdash_T \theta(\underline{c})$ and $\Theta_2 \vdash_T \neg\theta(\underline{c})$.*

*Then, there exist models $\mathcal{M}_1, \mathcal{M}_2$ of $T$ such that $\mathcal{M}_1 \models \Theta_1$, $\mathcal{M}_2 \models \Theta_2$ and the intersection of the supports of $\mathcal{M}_1$ and $\mathcal{M}_2$ coincides with the support of the substructure generated (in both of them) by the interpretation of the constants $\underline{c}$.*

[4] By this (and similar notions) we mean that $\phi \wedge \psi$ is unsatisfiable in all $\Sigma'$-structures whose $\Sigma$-reduct is a model of $T$.

[5] For the results of this paper to be correct, the notion of structure (and of course that of substructure) should encompass the case of structures with empty domains. Readers feeling uncomfortable with empty domains can assume that signatures always contain an individual constant.

PROOF. Let $\Sigma^A$ be the signature $\Sigma$ expanded with the free constants $\underline{a} \cup \underline{c}$ and $\Sigma^B$ the signature $\Sigma$ expanded with the free constants $\underline{b} \cup \underline{c}$ (we put $\Sigma^C := \Sigma^A \cap \Sigma^B = \Sigma \cup \{\underline{c}\}$). Below, we use $\Sigma^A, \Sigma^B, \Sigma^C$ to indicate not only the signatures $\Sigma^A, \Sigma^B, \Sigma^C$ but also the set of formulae in the signature $\Sigma^A, \Sigma^B, \Sigma^C$, respectively.

As a first step, we build a maximal $T$-consistent set $\Gamma$ of ground $\Sigma^A$-formulae and a maximal $T$-consistent set $\Delta$ of ground $\Sigma^B$-formulae such that $\Theta_1 \subseteq \Gamma$, $\Theta_2 \subseteq \Delta$, and $\Gamma \cap \Sigma^C = \Delta \cap \Sigma^C$. For simplicity[6] let us assume that $\Sigma$ is at most countable, so that we can fix two enumerations

$$\phi_1, \phi_2, \ldots \qquad \psi_1, \psi_2, \ldots$$

of ground $\Sigma^A$- and $\Sigma^B$-formulae, respectively. We build inductively $\Gamma_n, \Delta_n$ such that for every $n$ (i) $\Gamma_n$ contains either $\phi_n$ or $\neg\phi_n$; (ii) $\Delta_n$ contains either $\psi_n$ or $\neg\psi_n$; (iii) there is no ground $\Sigma^C$-formula $\theta$ such that $\Gamma_n \cup \{\neg\theta\}$ and $\Delta_n \cup \{\theta\}$ are not $T$-consistent. Once this is done, we can get our $\Gamma, \Delta$ as $\Gamma \equiv \bigcup \Gamma_n$ and $\Delta \equiv \bigcup \Delta_n$.

We let $\Gamma_0$ be $\Theta_1$ and $\Delta_0$ be $\Theta_2$ (notice that (iii) holds by assumption). To build $\Gamma_{n+1}$ we have two possibilities, namely $\Gamma_n \cup \{\phi_n\}$ and $\Gamma_n \cup \{\neg\phi_n\}$. Suppose they are both unsuitable because there are $\theta_1, \theta_2 \in \Sigma^C$ such that the sets

$$\Gamma_n \cup \{\phi_n, \neg\theta_1\}, \quad \Delta_n \cup \{\theta_1\}, \quad \Gamma_n \cup \{\neg\phi_n, \neg\theta_2\}, \quad \Delta_n \cup \{\theta_2\}$$

are all $T$-inconsistent. If we put $\theta \equiv \theta_1 \vee \theta_2$, we get that $\Gamma_n \cup \{\neg\theta\}$ and $\Delta_n \cup \{\theta\}$ are not $T$-consistent, contrary to induction hypothesis. A similar argument shows that we can also build $\Delta_n$.

Let now $\mathcal{M}_1$ be a model of $\Gamma$ and $\mathcal{M}_2$ be a model of $\Delta$. Consider the substructures $\mathcal{A}_1, \mathcal{A}_2$ of $\mathcal{M}_1, \mathcal{M}_2$ generated by the interpretations of the constants from $\Sigma^C$: since they satisfy the same literals from $\Sigma^C$ (because $\Gamma \cap \Sigma^C = \Delta \cap \Sigma^C$), we have that $\mathcal{A}_1$ and $\mathcal{A}_2$ are $\Sigma^C$-isomorphic. Up to renaming, we can suppose that $\mathcal{A}_1$ and $\mathcal{A}_2$ are just the same substructure. □

THEOREM 3.3. *A theory $T$ has the sub-amalgamation property iff it admits quantifier-free interpolants.*

PROOF. *Suppose first that $T$ has sub-amalgamation*; let $\phi, \psi$ be quantifier-free formulae such that $\phi \wedge \psi$ is not $T$-satisfiable. Let us replace variables with free constants in $\phi, \psi$; let us call $\Sigma^A$ the signature $\Sigma$ expanded with the free constants from $\phi$ and $\Sigma^B$ the signature $\Sigma$ expanded with the free constants from $\psi$ (we put $\Sigma^C := \Sigma^A \cap \Sigma^B$). For reduction, suppose that there is no ground formula $\theta$ such that: (a) $\phi$ $T$-entails $\theta$; (b) $\theta \wedge \psi$ is $T$-unsatisfiable; (c) only free constants from $\Sigma^C$ occur in $\theta$. By Lemma 3.2, taking $\Theta_1 := \{\phi\}, \Theta_2 := \{\psi\}$, we know that there are models $\mathcal{M}_1, \mathcal{M}_2$ of $T$ such that $\mathcal{M}_1 \models \phi, \mathcal{M}_2 \models \psi$ and such that the intersection of the supports of $\mathcal{M}_1$ and $\mathcal{M}_2$ is precisely the substructure generated by the interpretation of the constants from $\Sigma^C$ (let us we call this substructure $\mathcal{A}$ for short). By the sub-amalgamation property, there is a $T$-amalgam $\mathcal{M}$ of $\mathcal{M}_1$ and $\mathcal{M}_2$ over $\mathcal{A}$. Now $\phi, \psi$ are ground formulae true in $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, hence they are both true in $\mathcal{M}$, which is impossible because $\phi \wedge \psi$ was assumed to be $T$-inconsistent.

*Suppose now that $T$ has quantifier-free interpolants*. Take two models $\mathcal{M}_1$ and $\mathcal{M}_2$ of $T$ sharing a substructure $\mathcal{A}$; we can freely suppose (up to a renaming) that $|\mathcal{M}_1| \cap |\mathcal{M}_2| = |\mathcal{A}|$ (recall that we use the notation $|-|$ to indicated the support of a structure). In order to show that a $T$-amalgam of $\mathcal{M}_1, \mathcal{M}_2$ over $\mathcal{A}$ exists, it is sufficient (by Robinson Diagram Lemma) to show that $\Delta_\Sigma(\mathcal{M}_1) \cup \Delta_\Sigma(\mathcal{M}_2)$ is $T$-consistent, where (for $i = 1, 2$) $\Delta_\Sigma(\mathcal{M}_i)$ is the *diagram* of $\mathcal{M}_i$.

[6] This is just to avoid a (straightforward indeed) transfinite induction argument.

If $\Delta_\Sigma(\mathcal{M}_1) \cup \Delta_\Sigma(\mathcal{M}_2)$ is not $T$-consistent, by the compactness theorem of first order logic, there exist a $\Sigma \cup |\mathcal{M}_1|$-ground sentence $\phi$ and a $\Sigma \cup |\mathcal{M}_2|$-ground sentence $\psi$ such that (i) $\phi \wedge \psi$ is $T$-inconsistent; (ii) $\phi$ is a conjunction of literals from $\Delta_\Sigma(\mathcal{M}_1)$; (iii) $\psi$ is a conjunction of literals from $\Delta_\Sigma(\mathcal{M}_2)$. By the existence of quantifier-free interpolants, taking free constants instead of variables, we get that there exists a ground $\Sigma \cup |\mathcal{A}|$-sentence $\theta$ such that $\phi$ $T$-entails $\theta$ and $\psi \wedge \theta$ is $T$-inconsistent. The former fact yields that $\theta$ is true in $\mathcal{M}_1$ and hence also in $\mathcal{A}$ and in $\mathcal{M}_2$, because $\theta$ is ground. However, the fact that $\theta$ is true in $\mathcal{M}_2$ contradicts the fact that $\psi \wedge \theta$ is $T$-inconsistent. □

Theorem 3.3 allows us to derive in a uniform way results about quantifier-free interpolation that are known in the literature, as the following examples show.

*Example* 3.4. For any signature $\Sigma$, let $\mathcal{EUF}(\Sigma)$ be the pure theory of equality over $\Sigma$. This theory has quantifier-free interpolation (see, e.g., [McMillan 2004; Fuchs et al. 2009]). Indeed, it is easy to see that $\mathcal{EUF}(\Sigma)$ has the (strong) sub-amalgamation property by building a model $\mathcal{M}$ of $\mathcal{EUF}(\Sigma)$ from two models $\mathcal{M}_1$ and $\mathcal{M}_2$ sharing a substructure $\mathcal{A}$ as follows. Without loss of generality, assume that $|\mathcal{A}| = |\mathcal{M}_1| \cap |\mathcal{M}_2|$; let $|\mathcal{M}|$ be $|\mathcal{M}_1| \cup |\mathcal{M}_2|$ and arbitrarily extend the interpretation of the function and predicate symbols to make them total on $|\mathcal{M}|$.

*Example* 3.5. Theorem 3.3 can be extended in a straightforward way to many-sorted theories. Let us consider the two variants $\mathcal{AX}_{\mathrm{ext}}$ and $\mathcal{AX}_{\mathtt{diff}}$ of the theory of arrays introduced in [Bruttomesso et al. 2011b; 2012]. The signatures of $\mathcal{AX}_{\mathrm{ext}}$ and $\mathcal{AX}_{\mathtt{diff}}$ contain the sort symbols $\mathtt{ARRAY}, \mathtt{ELEM}$, and $\mathtt{INDEX}$, and the function symbols $rd : \mathtt{ARRAY} \times \mathtt{INDEX} \longrightarrow \mathtt{ELEM}$ and $wr : \mathtt{ARRAY} \times \mathtt{INDEX} \times \mathtt{ELEM} \longrightarrow \mathtt{ARRAY}$. The signature of $\mathcal{AX}_{\mathtt{diff}}$ also contains the function symbol $\mathtt{diff} : \mathtt{ARRAY} \times \mathtt{ARRAY} \longrightarrow \mathtt{INDEX}$. The set $\mathcal{AX}_{\mathrm{ext}}$ of axioms contains the following three sentences:

$$\forall y, i, j, e.\ i \neq j \Rightarrow rd(wr(y, i, e), j) = rd(y, j),$$
$$\forall y, i, e.\ rd(wr(y, i, e), i) = e,$$
$$\forall x, y.\ x \neq y \Rightarrow (\exists i.\ rd(x, i) \neq rd(y, i))$$

whereas the set of axioms for $\mathcal{AX}_{\mathtt{diff}}$ is obtained from that of $\mathcal{AX}_{\mathrm{ext}}$ by replacing the third axiom with its Skolemization:

$$\forall x, y. \quad x \neq y \Rightarrow rd(x, \mathtt{diff}(x, y)) \neq rd(y, \mathtt{diff}(x, y))\ .$$

In [Bruttomesso et al. 2012] (the journal version of [Bruttomesso et al. 2011b]), it is shown that $\mathcal{AX}_{\mathtt{diff}}$ has the sub-amalgamation property while $\mathcal{AX}_{\mathrm{ext}}$ does not ($\mathcal{AX}_{\mathtt{diff}}$ enjoys also the strong sub-amalgamation property). In fact, $\mathcal{AX}_{\mathtt{diff}}$ [Bruttomesso et al. 2011b; 2012] admits quantifier-free interpolants, whereas $\mathcal{AX}_{\mathrm{ext}}$ does not [Kapur et al. 2006]. Notice that $\mathcal{AX}_{\mathrm{ext}}$ (which is *not* universal) enjoys the following property (this is the standard notion of amalgamability from the literature): given two models $\mathcal{M}_1$ and $\mathcal{M}_2$ of $\mathcal{AX}_{\mathrm{ext}}$ sharing a substructure $\mathcal{M}_0$ *which is also a model of* $\mathcal{AX}_{\mathrm{ext}}$, there is a model $\mathcal{M}$ of $\mathcal{AX}_{\mathrm{ext}}$ endowed with embeddings from $\mathcal{M}_1, \mathcal{M}_2$ agreeing on the support of $\mathcal{M}_0$. This proves that amalgamation is not sufficient to guarantee quantifier-free interpolation property for non-universal theories; sub-amalgamation is required instead.

### 3.2. Modularity of quantifier-free interpolation

Given the importance of combining theories in SMT solving, the next step is to establish whether sub-amalgamation is a modular property. Unfortunately, the combination of two theories may not have quantifier-free interpolation despite the fact that each component theory does have it. To see this, consider, for example, the union of the theory of equality with uninterpreted function symbols (called $\mathcal{EUF}(\Sigma)$ in Example 3.4) and Presburger arithmetic: this theory does not have quantifier-free interpo-

lation [Brillout et al. 2010]. Fortunately, strong sub-amalgamation is modular when combining stably infinite theories. To prove it, we first recall the following lemma, which is an important ingredient of the correctness of the well-known Nelson-Oppen combination method [Nelson and Oppen 1979; Tinelli and Harandi 1996]; we state and prove the lemma in the form we directly use in the sequel:

LEMMA 3.6. *Suppose that $T_1, T_2$ are two stably infinite theories with disjoint signatures $\Sigma_1, \Sigma_2$ and let $C$ be a set of free constants not belonging to $\Sigma_1 \cup \Sigma_2$. Let $\Gamma$ be a partition of $C$, i.e. a set of ground equalities or inequalities containing the literal $c_1 = c_2$ or the literal $c_1 \neq c_2$, for all pairs of different constants from $C$. For $i = 1, 2$, let $\Theta_i$ be a $T_i$-consistent set of ground $\Sigma_i \cup C$-formulae containing $\Gamma$. Then, $\Theta_1 \cup \Theta_2$ is $T_1 \cup T_2$-consistent.*

PROOF. Let $\mathcal{M}_1, \mathcal{M}_2$ be two models of $T_1 \cup \Theta_1, T_2 \cup \Theta_2$, respectively. By stable infiniteness and upward Lövenheim-Skolem theorem [Chang and Keisler 1990], we can assume that they are both infinite and have the same cardinality (bigger than the cardinality of $C$). Thus there is a bijection $f$ among their supports and (as equalities of constants from $C$ are interpreted in the same way in $\mathcal{M}_1$ and $\mathcal{M}_2$) we can assume that $f(c^{\mathcal{M}_1}) = c^{\mathcal{M}_2}$ holds for all $c \in C$. Using this bijection, it is easy to lift the interpretation of the $\Sigma_2$-symbols from the support of $\mathcal{M}_2$ to the support of $\mathcal{M}_1$. The lifted model is $\Sigma_2 \cup C$-isomorphic to $\mathcal{M}_2$, thus it is a model of $T_1 \cup T_2 \cup \Theta_1 \cup \Theta_2$. □

Two sets of sentences $\Gamma$ and $\Delta$ are said to be *logically equivalent modulo a theory $T$* iff for every $\gamma \in \Gamma$ we have $\Delta \vdash_T \gamma$ and vice versa. They are said to be *logically equivalent tout court* iff they are logically equivalent modulo the empty theory.

LEMMA 3.7. *Let $\Sigma_1, \Sigma_2$ be two signatures and $\mathcal{A}$ be a $\Sigma_1 \cup \Sigma_2$-structure; then $\Delta_{\Sigma_1 \cup \Sigma_2}(\mathcal{A})$ is logically equivalent to $\Delta_{\Sigma_1}(\mathcal{A}) \cup \Delta_{\Sigma_2}(\mathcal{A})$.*

PROOF. As a general fact, let us first observe that every diagram is equivalent to its flat subdiagram, in the following sense. Let $\mathcal{B}$ be a $\Sigma$-structure; the flat $\Sigma$-subdiagram of $\mathcal{B}$ is the set $\Delta^f_\Sigma(\mathcal{B})$ of literals of the kind

$$f(a_1, \ldots, a_n) = b, \quad a_1 \neq a_2, \quad P(a_1, \ldots, a_n), \quad \neg P(a_1, \ldots, a_n)$$

that are true in $\mathcal{B}$ (here $a_1, \ldots, a_n, b$ are free constants naming elements from $|\mathcal{B}|$). That every diagram is equivalent to its flat subdiagram can be easily proved from the fact (to be shown preventively by induction) that for every $\Sigma \cup |\mathcal{B}|$-ground term $t$ there is $a \in |\mathcal{B}|$ such that $\Delta^f_\Sigma(\mathcal{B}) \vdash t = a$.

Now we have that $\Delta_{\Sigma_1 \cup \Sigma_2}(\mathcal{A})$ is logically equivalent to $\Delta^f_{\Sigma_1 \cup \Sigma_2}(\mathcal{A})$ and the latter is $\Delta^f_{\Sigma_1}(\mathcal{A}) \cup \Delta^f_{\Sigma_2}(\mathcal{A})$ which in turn is equivalent to $\Delta_{\Sigma_1}(\mathcal{A}) \cup \Delta_{\Sigma_2}(\mathcal{A})$. □

THEOREM 3.8. *Let $T_1$ and $T_2$ be two stably infinite theories over disjoint signatures $\Sigma_1$ and $\Sigma_2$. If both $T_1$ and $T_2$ have the strong sub-amalgamation property, then so does $T_1 \cup T_2$.*

PROOF. Consider two models $\mathcal{M}_1, \mathcal{M}_2$ of $T_1 \cup T_2$ together with a common substructure $\mathcal{A}$; we can freely suppose (up to a renaming) that $|\mathcal{M}_1| \cap |\mathcal{M}_2| = |\mathcal{A}|$. By Robinson Diagram Lemma [Chang and Keisler 1990] and Lemma 3.7, it is sufficient to show the consistency of $T_1 \cup T_2 \cup \Gamma_1 \cup \Gamma_2$, where $\Gamma_i$ ($i = 1, 2$) is defined as

$$\Gamma_i \;\equiv\; \Delta_{\Sigma_i}(\mathcal{M}_1) \cup \Delta_{\Sigma_i}(\mathcal{M}_2) \cup \{m_1 \neq m_2 \mid m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|,\ m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|\} \ .$$

By the strong amalgamability of $T_1$ and $T_2$, we know that $T_1 \cup \Gamma_1$ and $T_2 \cup \Gamma_2$ are both consistent. Now notice that for every pair $c_1, c_2$ of distinct constants, the set $\Gamma_i$ contains the negative literal $c_1 \neq c_2$: in fact, this inequation is part of the definition

of the diagram of a structure or (in case $c_1, c_2$ are from different supports) it has been added explicitly when building $\Gamma_1, \Gamma_2$. According to Lemma 3.6, this is sufficient to infer the consistency of $T_1 \cup T_2 \cup \Gamma_1 \cup \Gamma_2$, as $T_1, T_2$ are stably infinite. □

Theorems 3.3 and 3.8 obviously imply that strong sub-amalgamation is sufficient for the modularity of quantifier-free interpolation for stably infinite theories.

COROLLARY 3.9. *Let $T_1$ and $T_2$ be two stably infinite theories over disjoint signatures $\Sigma_1$ and $\Sigma_2$. If both $T_1$ and $T_2$ have the strong sub-amalgamation property, then $T_1 \cup T_2$ admits quantifier-free interpolation.*

We can also show that strong sub-amalgamation is a necessary condition, in a sense that is precisely characterized in the following result.

THEOREM 3.10. *Let $T$ be a theory admitting quantifier-free interpolation and $\Sigma$ be a signature disjoint from the signature of $T$ containing at least a unary predicate symbol. Then, $T \cup \mathcal{EUF}(\Sigma)$ has quantifier-free interpolation iff $T$ has the strong sub-amalgamation property.*

PROOF. (Below $\Sigma_T$ is the signature of $T$). Let $T$ be strongly sub-amalgamable and let $\mathcal{M}_1, \mathcal{M}_2$ be two models of $T \cup \mathcal{EUF}(\Sigma)$ sharing a substructure $\mathcal{A}$ (as usual, we suppose that $|\mathcal{M}_1| \cap |\mathcal{M}_2| = |\mathcal{A}|$). To amalgamate them, consider first a model $\mathcal{M}$ of $T$ strongly amalgamating the $\Sigma_T$-reducts of $\mathcal{M}_1, \mathcal{M}_2$ over the $\Sigma_T$-reduct of $\mathcal{A}$. Since the amalgam is strong, up to isomorphism we can consider the support of $\mathcal{M}$ as a superset of $|\mathcal{M}_1| \cup |\mathcal{M}_2|$; thus it is easy to expand $\mathcal{M}$ to a total structure interpreting the symbols of $\Sigma$. The expansion is a model of $T \cup \mathcal{EUF}(\Sigma)$ amalgamating $\mathcal{M}_1$ and $\mathcal{M}_2$ over $\mathcal{A}$. This shows that $T \cup \mathcal{EUF}(\Sigma)$ has the sub-amalgamation property and hence the quantifier-free interpolation property.

Conversely, suppose that $T$ does not have the strong sub-amalgamation property. Let $\mathcal{M}_1, \mathcal{M}_2$ be models of $T_1$ and let $\mathcal{A}$ be a substructure of them such that there is no triple $\mathcal{M}, \mu_1, \mu_2$ satisfying the conditions for the strong sub-amalgamability property. This means that the set

$$\Gamma \equiv \Delta_{\Sigma_T}(\mathcal{M}_1) \cup \Delta_{\Sigma_T}(\mathcal{M}_2) \cup \{m_1 \neq m_2 \mid m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|,\ m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|\}$$

is not $T$-consistent. By compactness, there are $m_1^1, \ldots m_1^k \in |\mathcal{M}_1| \setminus |\mathcal{A}|$ and $m_2^1, \ldots m_2^k \in |\mathcal{M}_2| \setminus |\mathcal{A}|$ such that

$$T \cup \Delta_{\Sigma_T}(\mathcal{M}_1) \cup \Delta_{\Sigma_T}(\mathcal{M}_2) \models \bigvee_{j=1}^{k} m_1^j = m_2^j \ . \tag{1}$$

Expand now $\mathcal{M}_1, \mathcal{M}_2$ to $\Sigma_T \cup \Sigma$-structures as follows: the $\Sigma$-symbols are interpreted arbitrarily (but in such a way that $\mathcal{A}$ remains a substructure of the expansions) apart from the unary predicate $P$, which is interpreted as the whole support of $\mathcal{M}_1$ in the expansion of $\mathcal{M}_1$ and as the support of $\mathcal{A}$ in the expansion of $\mathcal{M}_2$. From (1), it is then clear that sub-amalgamation (hence quantifier-free interpolation) fails for $T \cup \mathcal{EUF}(\Sigma)$: in fact, any $\mathcal{M} \models T$ amalgamating $\mathcal{M}_1, \mathcal{M}_2$ over $\mathcal{A}$, must identify some $m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|$ with some $m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|$, which is impossible as the interpretation of $P$ in $\mathcal{M}$ must agree with the interpretations of $P$ in the expansions of $\mathcal{M}_1$ and $\mathcal{M}_2$. □

The signature $\Sigma$ from the statement of Theorem 3.10 is assumed to contain at least a unary predicate symbol; of course, one may equivalently assume that $\Sigma$ contains an $n$-ary predicate symbol for some $n \geq 1$. It is not difficult to see that one can also equivalently assume that $\Sigma$ contains an $n$-ary function symbol (for some $n \geq 1$) in case there is at least an individual constant in the signature of $T$.

### 3.3. Interpolation with free function symbols

As already pointed out at the end of Section 2, the notion of quantifier-free interpolation studied in Sections 3.1 and 3.2 is different from that commonly adopted in verification, where free (function and predicate) symbols are not allowed to appear in the interpolants if they do not occur in the formulæ for which an interpolant is to be computed. This more general notion of interpolation is formally characterized in Definition 2.2; one may also argue that it is this general quantifier-free interpolation property that represents the 'quantifier-free modulo theory' version of the original interpolation property introduced by Craig for pure first order logic.

Interestingly, it is possible to show that the property of general quantifier-free interpolation for a theory $T$ can be reduced to the plain quantifier-free interpolation of the union of $T$ and the pure theory of equality with (a disjoint set of) uninterpreted symbols. In turn, this implies that the more general notion of quantifier-free interpolation used in verification is equivalent to strong sub-amalgamability.

Assume that the theory $T$ has the property of general quantifier-free interpolation of Definition 2.2. A natural question to ask is whether this is equivalent to require that the property of (plain) quantifier-free interpolation holds for the combined theory $T \cup \mathcal{EUF}(\Sigma')$, for any signature $\Sigma'$ that is disjoint from that of $T$. At a first, the answer seems to be negative as Definition 2.2 requires that also the function and the predicate symbols in $\Sigma'$ that occur neither in $\phi$ nor in $\psi$ should not occur in $\theta$. We shall see however that such symbols are immaterial—as they can be removed—and the answer to the question above is positive.

Let $T$ be a theory with a signature $\Sigma$ and $\Sigma'$ be a signature that is disjoint from $\Sigma$. A finite set $A$ of ground $\Sigma \cup \Sigma'$-formulæ is said to be $\Sigma_0$-*flat* (for some $\Sigma_0 \subseteq \Sigma'$) iff $A$ is of the form $A_0 \cup A_1$ where $A_1$ does not contain $\Sigma_0$-symbols and $A_0$ is a set of literals of the forms

$$f(a_1, \ldots, a_n) = b, \quad P(a_1, \ldots, a_n), \quad \neg P(a_1, \ldots, a_n),$$

$f, P \in \Sigma_0$, and $a_1, \ldots, a_n, b$ are constants not in $\Sigma_0$.

LEMMA 3.11. *Let $T, \Sigma, \Sigma'$ be as above and let the finite set of ground $\Sigma \cup \Sigma'$-formulæ $A$ be $\Sigma_0$-flat (for some $\Sigma_0 \subseteq \Sigma'$). Then it is possible to find a finite set of ground formulæ $A^{-\Sigma_0}$ such that: (i) $A^{-\Sigma_0}$ does not contain $\Sigma_0$-symbols; (ii) $A$ $T$-entails $A^{-\Sigma_0}$; (iii) $A^{-\Sigma_0}$ is $T$-satisfiable iff $A$ is $T$-satisfiable.*

PROOF. Let $A$ be $A_0 \cup A_1$ as prescribed in the definition of $\Sigma_0$-flatness. We take as $A^{-\Sigma_0}$ the set of ground formulæ $A'_0 \cup A_1$ where $A'_0$ is built as follows. For every function symbol $f \in \Sigma_0$ and for every pair of atoms $f(a_1, \ldots, a_n) = b, f(a'_1, \ldots, a'_n) = b'$ belonging to $A_0$ we include in $A'_0$ the ground clause

$$a_1 = a'_1 \wedge \cdots \wedge a_n = a'_n \rightarrow b = b'; \tag{2}$$

similarly, for every predicate symbol $P \in \Sigma_0$ and for every pair of literals $P(a_1, \ldots, a_n)$, $\neg P(a'_1, \ldots, a'_n)$ belonging to $A_0$ we include in $A'_0$ the ground clause

$$a_1 = a'_1 \wedge \cdots \wedge a_n = a'_n \rightarrow \bot. \tag{3}$$

That $A'_0 \cup A_1$ enjoys properties (i)-(ii) is clear; it remains to show that if it is $T$-satisfiable, so is $A_0 \cup A_1$ (the right-to-left side of (iii) is a consequence of (ii)). Suppose indeed that $\mathcal{M}$ is a $\Sigma \cup (\Sigma' \setminus \Sigma_0)$-model of $T$ in which $A'_0 \cup A_1$ is true. We expand $\mathcal{M}$ to a $\Sigma \cup \Sigma'$-structure as follows. Let $f \in \Sigma_0$ have arity $n$ and let $c_1, \ldots, c_n$ be elements from the support of $\mathcal{M}$; then $f^{\mathcal{M}}(c_1, \ldots, c_n)$ is arbitrary, unless there are $f(a_1, \ldots, a_n) = b \in A_0$ such that $c_1 = a_1^{\mathcal{M}}, \ldots, c_n = a_n^{\mathcal{M}}$: in this case, we put $f^{\mathcal{M}}(c_1, \ldots, c_n)$ to be equal to $b^{\mathcal{M}}$. Since $\mathcal{M}$ is a model of the clauses (2), the definition is correct. Similarly, if $P \in \Sigma_0$ has arity $n$, then $P^{\mathcal{M}}$ is the set of $n$-tuples $c_1, \ldots, c_n$

of elements from the support of $\mathcal{M}$ such that there exists $P(a_1, \ldots, a_n) \in A_0$ such that $c_1 = a_1^{\mathcal{M}}, \ldots, c_n = a_n^{\mathcal{M}}$. The literals from $A_0$ turns out to be all true by construction and because in $\mathcal{M}$ the clauses (3) hold. □

THEOREM 3.12. *$T$ has the general quantifier free interpolation property iff it is strongly sub-amalgamable.*

PROOF. Since the general quantifier-free interpolation property for $T$ implies the (plain) quantifier-free interpolation property for all theory $T \cup \mathcal{EUF}(\Sigma')$ for any $\Sigma'$ disjoint from the signature of $T$, it is clear from Theorem 3.10 that the general quantifier-free interpolation property implies strong sub-amalgamability. To show the converse, we use the Lemma 3.11 above together with the meta-rules introduced in our previous work [Bruttomesso et al. 2011b; 2012; 2011a]. To make the paper self-contained, the meta-rules and the related results will also be explained from scratch in Section 6.1 below, to which the reader is referred for details.

Let $\Sigma$ be the signature of $T$ and let $\Sigma'$ be disjoint from $\Sigma$; fix also finite sets of ground $\Sigma \cup \Sigma'$-formulae $A, B$ such that $A \cup B$ is $T$-unsatisfiable. Let $\Sigma_A$ be the set of predicate and (non-constant) function symbols from $\Sigma'$ that occur in $A$ but not in $B$; similarly, let $\Sigma_B$ be the set of predicate and (non-constant) function symbols from $\Sigma'$ that occur in $B$ but not in $A$. We show how to transform $A$ into a $\Sigma_A$-flat $\tilde{A}$ by using meta-rules (a similar transformation is applied to $B$ to get a $\Sigma_B$-flat $\tilde{B}$). Using meta-rules (Define1), (Redplus1), (Redminus1) (see Figure I in Subsection 6.1 below) we can add 'defining atoms' $f(a_1, \ldots, a_n) = a$ (with fresh $a$) and replace all occurrences of the term $f(a_1, \ldots, a_n)$ in $A$ by $a$; if we do it repeatedly, $A$ gets flattened, in the sense that function and predicate symbols (different from identity) in $A$ are always applied to constants. With this technique, we can transform $A$ into a conjunction of defining atoms and ground formulæ in which function symbols from $\Sigma_A$ do not occur. To take care of predicate symbols $P \in \Sigma_A$, we need guessing and meta-rule (Disjunction1): for every atom $P(a_1, \ldots, a_n)$ occurring in $A$, we add either $P(a_1, \ldots, a_n)$ or $\neg P(a_1, \ldots, a_n)$ to $A$ and replace $P(a_1, \ldots, a_n)$ with $\top$ or $\bot$, respectively (notice that because of such guessing the transformation from $A, B$ to $\tilde{A}, \tilde{B}$ may be non-deterministic). Since meta-rules are satisfiability-preserving and are endowed with recursive instructions for computation of interpolants (see Figure I and Proposition 6.1 below), it will be sufficient to find a desired interpolant $\theta$ for $\tilde{A}$ and $\tilde{B}$.

If we apply the transformations of Lemma 3.11 to $\tilde{A} \cup \tilde{B}$ we can get $(\tilde{A} \cup \tilde{B})^{-\Sigma_A} \equiv \tilde{A}^{-\Sigma_A} \cup \tilde{B}$ with the properties (i)-(iii) stated in that Lemma: in particular, function and predicate symbols from $\Sigma_A$ do not occur anymore in $\tilde{A}^{-\Sigma_A}$. We do the same for $\tilde{B}$ and eventually we get $\bar{A}, \bar{B}$ such that (a) $\bar{A} \cup \bar{B}$ is $T$-unsatisfiable; (b) $\tilde{A}$ $T$-entails $\bar{A}$, $\tilde{B}$ $T$-entails $\bar{B}$; (c) all predicate and (non-constant) functions symbols occurring in $\bar{A}$ occur also in $\bar{B}$ and vice versa. Let $\Sigma_C$ be the set of predicate and (non-constant) function symbols occurring in both $\bar{A}$ and $\bar{B}$. Since $T$ is strongly amalgamable, by Theorem 3.10, $T \cup \mathcal{EUF}(\Sigma_C)$ has the quantifier-free interpolation property.[7] Thus, there exists a ground formula $\theta$ containing, besides interpreted symbols from $\Sigma$, only predicate and function symbols from $\Sigma_C$, as well as individual free constants occurring both in $\bar{A}$ and in $\bar{B}$, such that $\bar{A}$ $T$-entails $\theta$ and $\bar{B} \wedge \theta$ is $T$-inconsistent. By (b) above, we get that $\tilde{A}$ $T$-entails $\theta$ and $\tilde{B} \wedge \theta$ is $T$-inconsistent, thus $\theta$ is the desired interpolant. □

[7] The proof of the right-to-left side of that Theorem does not need the requirement that $\Sigma_C$ has at least a unary predicate symbol.

## 4. EQUALITY INTERPOLATION AND STRONG AMALGAMATION

Although Corollary 3.9 is already useful to establish whether combinations of theories admit quantifier-free interpolants, proving the strong sub-amalgamability property can be complex. Below, we study an alternative ("operational") characterization of strong sub-amalgamability that can be more easily applied to theories that are considered important in applications.

There is a close relationship between the strong sub-amalgamation property of a theory $T$ and the fact that disjunctions of equalities among variables are entailed, modulo $T$, by finite sets of atoms. To make this precise, we need to introduce some preliminary notions.

Given two finite tuples $\underline{t} \equiv t_1, \ldots, t_n$ and $\underline{v} \equiv v_1, \ldots, v_m$ of terms,

$$\text{the notation } \underline{t} \cap \underline{v} \neq \emptyset \text{ stands for the formula } \bigvee_{i=1}^{n} \bigvee_{j=1}^{m} (t_i = v_j).$$

We use $\underline{t}_1\underline{t}_2$ to denote the juxtaposition of the two tuples $\underline{t}_1$ and $\underline{t}_2$ of terms. So, for example, $\underline{t}_1\underline{t}_2 \cap \underline{v} \neq \emptyset$ is equivalent to $(\underline{t}_1 \cap \underline{v} \neq \emptyset) \vee (\underline{t}_2 \cap \underline{v} \neq \emptyset)$.

*Definition* 4.1. A theory $T$ is *equality interpolating* iff it has the quantifier-free interpolation property and satisfies the following condition:

— for every quintuple $\underline{x}, \underline{y}_1, \underline{z}_1, \underline{y}_2, \underline{z}_2$ of tuples of variables and for every pair of quantifier-free formulae $\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1)$ and $\delta_2(\underline{x}, \underline{z}_2, \underline{y}_2)$ such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{y}_2 \neq \emptyset \tag{4}$$

there exists a tuple $\underline{v}(\underline{x})$ of terms such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \underline{y}_1\underline{y}_2 \cap \underline{v} \neq \emptyset\ . \tag{5}$$

We now formally state the equivalence between strong sub-amalgamability and equality interpolation.

THEOREM 4.2. *Given a theory $T$ having quantifier-free interpolation, the following conditions are equivalent:*

(i). *$T$ is strongly sub-amalgamable;*
(ii). *$T$ is equality interpolating;*
(iii). *for every triple $\underline{x}, \underline{y}_1, \underline{y}_2$ of tuples of variables and for every pair of quantifier-free formulae $\delta_1(\underline{x}, \underline{y}_1), \delta_2(\underline{x}, \underline{y}_2)$ such that*

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{y}_2 \neq \emptyset \tag{6}$$

*there exists a tuple $\underline{v}(\underline{x})$ of terms such that*

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1\underline{y}_2 \cap \underline{v} \neq \emptyset\ . \tag{7}$$

PROOF. We first prove (i) $\Rightarrow$ (ii). Suppose that $T$ is strongly sub-amalgamable, we show that (4) $\Rightarrow$ (5) holds by contraposition. For this, let us fix the tuples of fresh free constants $\underline{a}, \underline{m}_1, \underline{n}_1, \underline{m}_2, \underline{n}_2$ and assume that for every finite tuple $\underline{v}$ of $\Sigma \cup \{\underline{a}\}$-ground terms, the formula

$$\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{n}_2, \underline{m}_2) \wedge (\underline{m}_1\underline{m}_2 \cap \underline{v} = \emptyset) \tag{8}$$

is $T$-consistent, where $\Sigma$ is the signature of $T$. We claim that the set

$$\{\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1), \delta_2(\underline{a}, \underline{n}_2, \underline{m}_2)\} \cup \{\underline{m}_1\underline{m}_2 \cap \underline{v} = \emptyset\}_{\underline{v}} \tag{9}$$

is $T$-consistent, where $\underline{v}$ varies over all possible tuples of such terms. In fact, if (9) were not consistent, by compactness, there would be tuples of $\Sigma \cup \{\underline{a}\}$-ground terms $\underline{v}_1, \ldots, \underline{v}_k$ such that

$$\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{n}_2, \underline{m}_2) \wedge \bigwedge_{j=1}^{k} (\underline{m}_1 \underline{m}_2 \cap \underline{v}_j = \emptyset)$$

were not $T$-consistent. Putting $\underline{v}$ equal to the tuple obtained by juxtaposition $\underline{v}_1 \cdots \underline{v}_k$, we would get a $\underline{v}$ contradicting (8).

Let $\Theta_1$ be $\{\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1)\} \cup \{\underline{m}_1 \cap \underline{v} = \emptyset\}_{\underline{v}}$ and let $\Theta_2$ be $\{\delta_2(\underline{a}, \underline{n}_2, \underline{m}_2)\} \cup \{\underline{m}_2 \cap \underline{v} = \emptyset\}_{\underline{v}}$. Since $\Theta_1 \cup \Theta_2$ is equal to (9) which is $T$-consistent, there is no ground $\Sigma \cup \{\underline{a}\}$-formula $\theta(\underline{a})$ such that $\Theta_1 \vdash_T \theta(\underline{a})$ and such that $\Theta_2 \cup \{\theta(\underline{a})\}$ is not $T$-consistent. By Lemma 3.2, we can then produce models $\mathcal{M}_1, \mathcal{M}_2$ of $T$ such that $\mathcal{M}_1 \models \Theta_1$, $\mathcal{M}_2 \models \Theta_2$ and such that the intersection of their supports is precisely the substructure generated by the interpretation of the constants $\underline{a}$. If we now strongly amalgamate them, we get a model of $T$ in which $\delta_1(\underline{a}, \underline{n}_1, \underline{m}_1), \delta_2(\underline{a}, \underline{n}_2, \underline{m}_2), \underline{m}_1 \cap \underline{m}_2 = \emptyset$ are all true, showing that (4) fails.

Since (ii) $\Rightarrow$ (iii) is trivial, we consider the proof of the implication (iii) $\Rightarrow$ (i).[8] Suppose that (6) $\Rightarrow$ (7) holds: we prove strong sub-amalgamability. If the latter property fails, by Robinson Diagram Lemma, there exist models $\mathcal{M}_1, \mathcal{M}_2$ of $T$ together with a shared substructure $\mathcal{A}$ such that the set of sentences

$$\Gamma \;\equiv\; \Delta_\Sigma(\mathcal{M}_1) \cup \Delta_\Sigma(\mathcal{M}_2) \cup \{m_1 \neq m_2 \mid m_1 \in |\mathcal{M}_1| \setminus |\mathcal{A}|,\; m_2 \in |\mathcal{M}_2| \setminus |\mathcal{A}|\}$$

is not $T$-consistent. By compactness, the sentence

$$\delta_1(\underline{a}, \underline{m}_1) \wedge \delta_2(\underline{a}, \underline{m}_2) \rightarrow \underline{m}_1 \cap \underline{m}_2 \neq \emptyset$$

is $T$-valid, for some tuples $\underline{a} \subseteq |\mathcal{A}|$, $\underline{m}_1 \subseteq (|\mathcal{M}_1| \setminus |\mathcal{A}|)$, $\underline{m}_2 \subseteq (|\mathcal{M}_2| \setminus |\mathcal{A}|)$ and for some ground formulae $\delta_1(\underline{a}, \underline{m}_1), \delta_2(\underline{a}, \underline{m}_2)$ true in $\mathcal{M}_1, \mathcal{M}_2$, respectively. By the implication (6) $\Rightarrow$ (7), there exists a finite tuple $\underline{v}(\underline{a})$ of $\Sigma \cup \{\underline{a}\}$-terms such that

$$\delta_1(\underline{a}, \underline{m}_1) \wedge (\underline{m}_1 \cap \underline{v}(\underline{a}) = \emptyset) \wedge \delta_2(\underline{a}, \underline{m}_2) \wedge (\underline{m}_2 \cap \underline{v}(\underline{a}) = \emptyset)$$

is not $T$-consistent. Since $T$ has quantifier-free interpolation, there is a ground formula $\theta(\underline{a})$ such that

$$\delta_1(\underline{a}, \underline{m}_1) \wedge (\underline{m}_1 \cap \underline{v}(\underline{a}) = \emptyset) \rightarrow \theta(\underline{a}) \tag{10}$$

is $T$-valid and

$$\delta_2(\underline{a}, \underline{m}_2) \wedge (\underline{m}_2 \cap \underline{v}(\underline{a}) = \emptyset) \wedge \theta(\underline{a}) \tag{11}$$

is not $T$-consistent. This is a contradiction, since $\underline{m}_1 \subseteq |\mathcal{M}_1| \setminus |\mathcal{A}|$ and the formula $\underline{m}_1 \cap \underline{v}(\underline{a}) = \emptyset$ is true in $\mathcal{M}_1$, which entails that $\theta(\underline{a})$ is true in $\mathcal{A}$ and in $\mathcal{M}_2$ too, where (11) consequently holds. □

Notice that (iii) is a special case of (ii) in which the tuple $\underline{z}$ is empty. It turns out that this is enough for several applications of the theorem; e.g., the combined interpolation algorithm in Section 6 below exploits only condition (iii).

### 4.1. Equality interpolation and quantifier elimination

We now illustrate some interesting applications of Theorem 4.2 that, with Corollary 3.9, allow us to establish when combinations of theories admit quantifier-free interpolation.

[8] Only in the proof of this implication we need the hypothesis that $T$ has quantifier-free interpolation.

The key notion is that of quantifier elimination (see, e.g., [Enderton 1972]): it is easy to see that quantifier elimination implies quantifier-free interpolation. In fact, the Craig interpolation theorem in first order logic ensures the existence of an interpolant and quantifier elimination ensures that a quantifier-free formula, that is logically equivalent to the interpolant, can always be obtained. The converse is not true:[9] $\mathcal{EUF}$ has quantifier-free interpolation but it does not enjoy quantifier elimination; otherwise first order logic would be decidable, since satisfiability of quantifier-free formulæ is decidable!

The fact that a theory enjoys quantifier elimination is not sufficient, alone, to guarantee that the theory is equality interpolating or, equivalently, strongly amalgamable; Presburger arithmetic being a counterexample (see below). For theories admitting quantifier elimination that have a set of *universal* axioms, the situation is different.

LEMMA 4.3. *Let $T$ be a theory admitting quantifier elimination. Then, $T$ is universal iff for every quantifier-free formula $\phi(\underline{x}, \underline{y})$, there exist tuples $\underline{t}_1(\underline{x}), \ldots, \underline{t}_n(\underline{x})$ of tuples of terms such that*

$$T \vdash \exists \underline{y}\, \phi(\underline{x}, \underline{y}) \leftrightarrow \bigvee_{i=1}^{n} \phi(\underline{x}, \underline{t}_i(\underline{x})) \quad . \tag{12}$$

PROOF. If the condition of the Lemma is true for every $\phi(\underline{x}, \underline{y})$, one can find an equivalent universal set of axioms for $T$ as follows. Notice that the right-to-left direction of (12) is a logical validity and the left-to-right direction is equivalent to a universal formula. Thus, we can take as axioms of $T$ the universal closures of the left-to-right directions of (12), together with the ground formulae which are logical consequences of $T$. In fact, axioms (12) are sufficient to find, for every sentence, a ground formula $T$-equivalent to it.

Conversely, suppose that $T$ is universal and that there is $\phi(\underline{x}, y_1, \ldots, y_m)$ such that (12) does not hold (for all possible tuples of $m$-tuples of terms). Then, by compactness, we have that the set of sentences

$$\Gamma \;\equiv\; \{\phi(\underline{a}, \underline{b})\} \cup \{\neg\phi(\underline{a}, \underline{t}(\underline{a}))\}_{\underline{t}}$$

is $T$-consistent (here $\Sigma$ is the signature of $T$, $\underline{a}, \underline{b} := b_1, \ldots, b_m$ are tuples of fresh constants and $\underline{t}$ vary on the set of $m$-tuples of $\Sigma \cup \{\underline{a}\}$-terms). Let $\mathcal{M}$ be a $T$-model of $\Gamma$ and let $\mathcal{N}$ be the substructure of $\mathcal{M}$ generated by the $\underline{a}$. Since $T$ is universal and truth of universal sentences is preserved under taking substructures, $\mathcal{N}$ is also a model of $T$ and since $T$ has quantifier-elimination, $\exists \underline{y}\, \phi(\underline{a}, \underline{y})$—being $T$-equivalent to a quantifier-free $\Sigma \cup \{\underline{a}\}$-sentence— is also true in $\mathcal{N}$. This is a contradiction because from $\mathcal{N} \models \exists \underline{y}\, \phi(\underline{a}, \underline{y})$ it follows that $\mathcal{N} \models \phi(\underline{a}, \underline{t}(\underline{a}))$ holds for some $\underline{t}$, contrary to the fact that $\mathcal{M} \not\models \phi(\underline{a}, \underline{t}(\underline{a}))$ and to the fact that $\mathcal{N}$ is a substructure of $\mathcal{M}$. □

THEOREM 4.4. *A universal theory admitting quantifier elimination is equality interpolating.*

PROOF. We show that a universal and quantifier eliminable theory $T$ satisfies the implication (6) $\Rightarrow$ (7). Suppose that (6) holds; by the previous Lemma, there exist tuples of terms $\underline{t}_1(\underline{x}), \ldots, \underline{t}_k(\underline{x})$ such that

$$\exists \underline{y}_2\, \delta_2(\underline{x}, \underline{y}_2) \leftrightarrow \bigvee_{j=1}^{k} \delta_2(\underline{x}, \underline{t}_j(\underline{x})) \tag{13}$$

[9] This is not in contrast with the results in [Kapur et al. 2006], because—in that paper—the notion of quantifier-free interpolation is formulated in a different (and much stronger) way by requiring that *all* pairs of mutually inconsistent formulæ—not just those without quantifiers—have a quantifier-free interpolant.

is $T$-valid. For every $j = 1, \ldots, k$, if we replace $\underline{y}_2$ with $\underline{t}_j$ in (6), we get

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{t}_j) \vdash_T \underline{y}_1 \cap \underline{t}_j \neq \emptyset$$

hence also

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \bigvee_{j=1}^{k} \delta_2(\underline{x}, \underline{t}_j) \vdash_T \bigvee_{j=1}^{k} (\underline{y}_1 \cap \underline{t}_j \neq \emptyset) \ .$$

Taking into account (13) and letting $\underline{v}$ be the tuple $\underline{t}_1 \cdots \underline{t}_k$ obtained by juxtaposition, we get

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \exists \underline{y}_2 \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{v} \neq \emptyset \ .$$

Removing the existential quantifier in the antecedent of the implication, we obtain

$$\delta_1(\underline{x}, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{y}_2) \vdash_T \underline{y}_1 \cap \underline{v} \neq \emptyset$$

and a fortiori (7), as desired. □

The obvious converse of Theorem 4.4 does not hold; in fact, the theory of dense linear orders without endpoints (see, e.g., [Chang and Keisler 1990]) has quantifier elimination and is equality interpolating (as it has the strong sub-amalgamation property) but it does not admit a universal set of axioms (since it is not closed under substructures).

We now consider some interesting theories to which Theorem 4.4 can be applied.

*Example* 4.5. The theory $\mathcal{RDS}$ of *recursive data structures* [Oppen 1980] consists of two unary function symbols $car$ and $cdr$ and a binary function symbol $cons$, and it is axiomatized by the following infinite set of sentences:

$$\forall x, y. car(cons(x, y)) = x, \tag{14}$$
$$\forall x, y. cdr(cons(x, y)) = y, \tag{15}$$
$$\forall x, y. cons(car(x), cdr(x)) = x, \tag{16}$$
$$\forall x. x \neq t(x) \tag{17}$$

where $t$ is a term obtained by finitely many applications of $car$ and $cdr$ to the variable $x$ (e.g., $car(x) \neq x$, $cdr(cdr(x)) \neq x$, $cdr(car(x)) \neq x$, and so on). Clearly, $\mathcal{RDS}$ is universal; the fact that it admits elimination of quantifiers is known since an old work by Mal'cev [Mal'cev 1962].

*Example* 4.6. Following [Enderton 1972], we define the theory $\mathcal{IDL}$ of *integer difference logic* to be the theory whose signature contains the constant symbol $0$, the unary function symbols $succ$ and $pred$, and the binary predicate symbol $<$, and which is axiomatized by adding to the irreflexivity, transitivity and linearity axioms for $<$ the following set of sentences:

$$\forall x. succ(pred(x)) = x, \qquad \forall x. pred(succ(x)) = x,$$
$$\forall x, y. x < succ(y) \leftrightarrow (x < y \vee x = y), \qquad \forall x, y. pred(x) < y \leftrightarrow (x < y \vee x = y).$$

$\mathcal{IDL}$ is universal and the fact that admits elimination of quantifiers can be shown by adapting the procedure for a similar theory of natural numbers with successor and ordering in [Enderton 1972]. The crucial observation is that the atoms of $\mathcal{IDL}$ are equivalent to formulae of the form $i \bowtie f^n(j)$ (for $n \in \mathbb{Z}$, $\bowtie \in \{=, <\}$) where $i, j$ are variables or the constant $0$, $f^0(j)$ is $j$, $f^k(j)$ abbreviates $succ(succ^{k-1}(j))$ when $k > 0$ or $pred(pred^{k-1}(j))$ when $k < 0$. Usually, $i \bowtie f^n(j)$ is written as $i - j \bowtie n$ or as $i \bowtie j + n$ from which the name of "integer difference logic."

*Example* 4.7. Presburger Arithmetic $\mathcal{PRA}$ can be specified as follows. The signature of $\mathcal{PRA}$ contains the symbols $0, 1, -, +, <$ (the first two are constants, the third is a unary function, the fourth is a binary function, and the last one is a binary predicate written infix) and in addition infinitely many unary predicates $P_n$, one for every $n > 0$. A set of axioms of $\mathcal{PRA}$ is the following (taken from [Dries ]):

$$\begin{array}{l}
\forall x, y, z.\ x + (y + z) = (x + y) + z \\
\forall x, y.\ x + y = y + x \\
\forall x.\ x + 0 = x \\
\forall x.\ x + (-x) = 0 \\
\forall x.\ x \not< x \\
\forall x, y, z.\ (x < y \wedge y < z \rightarrow x < z) \\
\forall x, y.\ x < y \vee x = y \vee y < x \\
\forall x, y, z.\ x < y \rightarrow x + z < y + z, \\
0 < 1, \\
\forall y. \neg(0 < y \wedge y < 1), \\
\forall x\, \exists y.\ \bigvee_{0 \leq r < n} x = ny + r \\
\forall x.\ P_n(x) \leftrightarrow \exists y\ (ny = x)
\end{array}$$

where $nt$ is an abbreviation for the sum of $n$-copies of $t$ and $n$ for the sum of $n$-copies of $1$. Notice that the last two axioms are in fact axioms schemata, representing infinitely many formulae. Presburger arithmetic enjoys quantifier-elimination: detailed proofs can be found in, e.g., [Enderton 1972; Dries ]. However, $\mathcal{PRA}$ is not equality interpolating as $\mathcal{PRA} \cup \mathcal{EUF}$ does not have the quantifier-free interpolation property [Brillout et al. 2010].

To overcome this problem, we suggest to modify $\mathcal{PRA}$ and derive an equivalent theory $\mathcal{LIA}$ as follows. First, we add the unary function symbols $div[n]$ (integer division by $n$), for $n > 1$. We then remove the last two axioms from the above list and replace them with the following one:

$$\forall x.\ \ x\ rem[n] = 0 \vee \cdots \vee x\ rem[n] = n - 1\,, \tag{18}$$

where $x\ rem[n]$ abbreviates $x - n(x\ div[n])$.

The crucial observation is that, in $\mathcal{LIA}$, $P_n(x)$ can be defined as $x\ rem[n] = 0$. Using this definition, we can view $\mathcal{LIA}$ as a supertheory of $\mathcal{PRA}$, because the last two axioms of $\mathcal{PRA}$ can be derived from those of $\mathcal{LIA}$.[10]

While $\mathcal{PRA}$ is not universal and thus does not allow the application of Theorem 4.4, $\mathcal{LIA}$ is so. Thus, if we are able to show that $\mathcal{LIA}$ admits quantifier elimination as it was the case for $\mathcal{PRA}$, we are entitled to conclude that it is equality interpolating (by Theorem 4.4). Interestingly, it is possible to reuse the available quantifier elimination procedures for $\mathcal{PRA}$ by pre-processing the formulae of $\mathcal{LIA}$ as explained in the proof of the following result.

*Proposition* 4.8. $\mathcal{LIA}$ has elimination of quantifiers and hence, being universal, it is equality interpolating.

PROOF. Let $\phi(\underline{x})$ be an arbitrary formula of $\mathcal{LIA}$; consider an atom $L$ occurring in $\phi$ containing an occurrence of a term $u$ of the form $t\ div[n]$. The atom $L$ is equivalent

[10] For the last one, show that the following universal sentences are derivable in $\mathcal{LIA}$ for every $n > 0$:

$$\forall x.\ nx = 0 \rightarrow x = 0 \qquad \forall x.\ \bigwedge_{0 < r < n} nx \neq r\ .$$

(modulo $\mathcal{LIA}$) to

$$\exists y \bigvee_{0 \leq r < n} (t = ny + r \wedge L[y/u]) \,. \tag{19}$$

This is because $\bigvee_{0 \leq r < n}(t = ny + r) \leftrightarrow y = t \; div[n]$ follows from the axioms of $\mathcal{LIA}$. We can then replace $L$ by (19) in $\phi$ and get an equivalent formula. If we do this exhaustively, we obtain a formula $\phi'$ such that $\mathcal{LIA} \vdash \phi \leftrightarrow \phi'$. Since, as we observed above, $\mathcal{LIA}$ is a supertheory of $\mathcal{PRA}$ and the latter enjoys quantifier elimination, we can find a quantifier-free $\phi''(\underline{x})$ such that $\mathcal{LIA} \vdash \phi \leftrightarrow \phi''$. □

Theorem 4.4 and Lemma 4.3 are not only useful when we are able to find a universal theory admitting elimination of quantifiers. In fact, Lemma 4.3 gives the possibility of checking the existence of a set of universal axioms just by inspecting the quantifier elimination procedure without the need to explicitly list its elements. The following two examples show this point.

*Example* 4.9. Consider Linear Arithmetic over the Reals. It is not difficult to see that the Fourier-Motzkin algorithm for quantifier elimination (see, e.g., [Dantzig and Eaves 1973]) satisfies the condition of Lemma 4.3, in the sense that it always eliminates an existential quantifier via a disjunction of instances.

For instance, when eliminating $\exists x$ from $\exists x\,(x < y_1 \wedge x < y_2 \wedge y_3 < x)$ one gets

$$(t_1 < y_1 \wedge t_1 < y_2 \wedge y_3 < t_1) \vee (t_2 < y_1 \wedge t_2 < y_2 \wedge y_3 < t_2)$$

where $t_1 := y_3 + (y_1 - y_3)/2$ and $t_2 := y_3 + (y_2 - y_3)/2$.

The example can be generalized and we can say that the Fourier-Motzkin algorithm eliminates an existentially quantified variable by substitutions with terms. This observation allows us to conclude that Linear Arithmetic over the reals has a universal set of axioms (it is not difficult to identify such axioms).

Notice that, for the above argument to work and for the axioms to be universal, is essential to include *multiplication by rational coefficients* in the signature of the theory, i.e. one needs (besides $+, -, 0, <$) the unary function symbols $q*\_$ for every $q \in \mathbb{Q}$. If this is not the case, the theory is not sub-amalgamable and thus not equality interpolating. To find a counterexample to strong amalgamation, consider the embedding of the substructure $\mathbb{Z}$ into two copies of the reals ($\mathbb{Z}$ is in fact a substructure of $\mathbb{R}$ if $+, -, 0, <$ are the only symbols in the language). If multiplication by rational coefficients is not in the language, one can get also a direct syntactic counterexample to the implication $(6) \Rightarrow (7)$ of Theorem 4.2(iii). To simplify the counterexample, since Linear Arithmetic over the Reals is convex, we can equivalently produce a counterexample to the implication $(20) \Rightarrow (21)$ of Proposition 5.2 (see below). To produce the counterexample, take $\delta_i(x, y_i) \equiv y_i + y_i = x$ for $i = 1, 2$: to get (21), one must use $v(x) \equiv \frac{1}{2} * x$ and this shows that the function symbol $\frac{1}{2} * \_$ is required to be in the language.

*Example* 4.10. The theory $\mathcal{UTVPI}$ is a fragment of Linear Arithmetic over the integers that is slightly more expressive than $\mathcal{IDL}$ (introduced in Example 4.6). It can be defined as the theory whose axioms are the sentences true in $\mathbb{Z}$ in the signature comprising the constant $0$, the unary function symbols $pred$, $succ$, and $-$, the binary predicate symbol $<$ (usually written infix). We shall prove here the existence of a set of universal quantifier eliminating axioms for $\mathcal{UTVPI}$, thus showing that $\mathcal{UTVPI}$ is also equality interpolating because of Theorem 4.4.

Preliminarily, we consider the atoms of $\mathcal{UTVPI}$ as we have done for $\mathcal{IDL}$. Such atoms are equivalent to expressions of the form $\pm i \bowtie f^n(j)$ (for $n \in \mathbb{Z}$, $\bowtie \in \{=, <, >\}$)[11]

[11] We use $>$ as a defined symbol: $t > u$ stands for $u < t$.

where $i, j$ are variables or the constant $0$, $f^0(j)$ is $j$, $f^k(j)$ abbreviates $succ(succ^{k-1}(j))$ when $k > 0$ or $pred(pred^{k-1}(j))$ when $k < 0$. Usually, $\pm i \bowtie f^n(j)$ is written as $i \pm j \bowtie n$ or as $i \bowtie n \pm j$.

*Proposition* 4.11. $\mathcal{UTVPI}$ is equality interpolating.

PROOF. Since we want to apply Lemma 4.3, we design a quantifier elimination algorithm for $\mathcal{UTVPI}$ satisfying (12). The algorithm guarantees the existence of a universal axiomatization for the theory—because of Lemma 4.3—and its being equality interpolating—by Theorem 4.4.

As usual when describing quantifier elimination procedures (see, e.g., [Chang and Keisler 1990; Enderton 1972]), it is sufficient to eliminate single existentially quantified variables from primitive formulæ. This means that, since negation can be eliminated, it is sufficient to consider formulae of the form $\exists x\, \phi$ where $\phi$ is a conjunction of atoms of the following types:

$$x = m_i \pm t_i, \qquad x < m_j \pm u_j, \qquad x > m_k \pm v_k,$$

where $x$ does not occur in the $t_i, u_j, v_k$ (otherwise either $\phi$ is inconsistent or the atom is redundant or it simplifies to an atom of the above kinds).

If there are literals of the first type, the quantifier $\exists x$ can be eliminated by substitution (this schema fits (12)), so suppose there are none. If there are no literals of the second type or no literals of the third, the formula $\exists x\, \phi$ is equivalent to $\top$ (use the terms $pred(m_j \pm u_j), succ(m_k \pm v_k)$ to fit (12)). If there are both literals of the second and of the third type, the formula $\exists x\, \phi$ is equivalent to $\bigvee_k \phi(succ(m_k \pm v_k))$. □

We conclude this part by observing that we will see another application of Theorem 4.4 and Lemma 4.3 in Section 6 when designing the interpolation algorithm for combination of theories. Such an algorithm takes as input formulæ $\delta_1, \delta_2$ satisfying (6) and is asked to compute terms $\underline{v}(\underline{x})$ satisfying (7). When the equality interpolating theory is universal and has quantifier elimination, one way to do this is to run the quantifier elimination algorithm over $\exists \underline{y_2}\, \delta_2(\underline{x}, \underline{y_2})$ and define $\underline{v}$ to be the tuple $\underline{t}_1 \cdots \underline{t}_k$ obtained by juxtaposition from the tuples in the right member of (13).

## 5. COMPARISON WITH PREVIOUS WORK

In the literature, there are two notions of equality interpolating theories, both introduced by Yorsh and Musuvathi. The former [Yorsh and Musuvathi 2005] is designed for convex theories and the latter—published only in the Technical Report [Yorsh and Musuvathi 2004] accompanying [Yorsh and Musuvathi 2005]—for non-convex theories.

Here, we first relate our notion of equality interpolating theory to those in [Yorsh and Musuvathi 2005; 2004] for convex (Section 5.1) and non-convex (Section 5.2) theories. Then, we show that notion of equality interpolation for convex theories of [Yorsh and Musuvathi 2005] is not sufficient to guarantee the modularity of quantifier-free interpolation for non-convex theories (Section 5.3). For the sake of completeness, we conclude by relating the notions introduced in this paper concerning quantifier-free interpolation and its modularity with standard notions in mathematical logic and universal algebra (Section 5.4).

### 5.1. The convex case

We show that our notion of equality interpolation reduces to that in [Yorsh and Musuvathi 2005] when considering convex theories only. We recall the formal definition of [Yorsh and Musuvathi 2005].

*Definition* 5.1. A theory $T$ has the *YMc property* (read as the 'convex Yorsh-Musuvathi property') iff it has the quantifier-free interpolation property and satisfies the following condition:

— for every pair $y_1, y_2$ of variables and for every pair of *conjunctions of literals* $\delta_1(\underline{x}, \underline{z}_1, y_1), \delta_2(\underline{x}, \underline{z}_2, y_2)$ such that

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2 \tag{20}$$

there exists a term $v(\underline{x})$ such that

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v \wedge y_2 = v. \tag{21}$$

PROPOSITION 5.2. *A convex theory $T$ is equality interpolating iff it has the YMc property.*

PROOF. If $\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2$ holds and $T$ is equality interpolating, it follows that there are terms $\underline{v}(\underline{x}) := v_1(\underline{x}), \ldots, v_n(\underline{x})$ such that

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T \bigvee_{i=1}^{n} (y_1 = v_i) \vee \bigvee_{i=1}^{n} (y_2 = v_i). \tag{22}$$

Let $w_1, \ldots, w_n$ be fresh variables; from (22) it follows that

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \wedge \bigwedge_{i=1}^{n} (w_i = v_i) \vdash_T \bigvee_{i=1}^{n} (y_1 = w_i) \vee \bigvee_{i=1}^{n} (y_2 = w_i).$$

Applying convexity, we obtain that there is some $i$ such that either

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \wedge \bigwedge_{i=1}^{n} (w_i = v_i) \vdash_T y_1 = w_i$$

or

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \wedge \bigwedge_{i=1}^{n} (w_i = v_i) \vdash_T y_2 = w_i$$

holds. Replacing the $w$'s with the $v$'s, this gives either

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v_i$$

or

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_2 = v_i.$$

In both cases (taking into consideration (20)), we get $\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = v_i \wedge y_2 = v_i$, as required by (21).

Vice versa, assume that the implication (20) $\Rightarrow$ (21) holds (for conjunctions of literals, as stated in Definition 5.1) and pick quantifier-free formulæ $\delta_1, \delta_2$ such that (4) holds. Let

$$\delta_1 \equiv \bigvee_j \theta_{1j}, \qquad \delta_2 \equiv \bigvee_k \theta_{2k} \ ,$$

where the $\theta_{ij}$ are conjunctions of literals. Applying convexity together with the implication (20) $\Rightarrow$ (21), we can find (single) terms $v_{jk}(\underline{x})$ satisfying $\theta_{1j} \wedge \theta_{2k} \vdash_T \underline{y}_1 \underline{y}_2 \cap v_{jk} \neq \emptyset$; if we let $\underline{v}$ be the tuple formed by all such $v_{jk}$, we finally get that the tuple of terms $\underline{v}$ satisfies (5). [12] $\square$

[12] This argument can be slightly modified to show that in Definition 4.1, we can restrict $\delta_1, \delta_2$ to be conjunctions of literals, getting anyway an equivalent definition.

Proposition 5.2 states the equivalence between the YMc property and that stated in Definition 4.1 for the class of convex theories. As a consequence, all the theories shown to have the YMc property in [Yorsh and Musuvathi 2005] are also equality interpolating in our sense. For example, consider the theory $\mathcal{LST}$ of list structures [Nelson and Oppen 1979]. Its signature contains the function symbols of $\mathcal{RDS}$ (recall Example 4.5) plus the unary predicate symbol $atom$. The set of axioms of $\mathcal{LST}$ contains the sentences (14) and (15) of $\mathcal{RDS}$ (see, again, Example 4.5) and

$$\forall x, y. \neg atom(cons(x, y)),$$
$$\forall x. \neg atom(x) \rightarrow cons(car(x), cdr(x)) = x\,.$$

$\mathcal{LST}$ is a (universal) convex theory [Nelson and Oppen 1979] that was shown to have the YMc property in [Yorsh and Musuvathi 2005]. By Proposition 5.2, we conclude that $\mathcal{LST}$ is equality interpolating in the sense of Definition 4.1. In [Yorsh and Musuvathi 2005], also Linear Arithmetic over the Reals ($\mathcal{LAR}$) is shown to enjoy the YMc property. By Proposition 5.2, $\mathcal{LAR}$ is equality interpolating in the sense of Definition 4.1. As we have seen in Example 4.9, the same result can be obtained from Theorem 4.4 by identifying a set of universal axioms for the theory and showing that quantifier elimination holds. The situation is similar for the pure theory of equality with uninterpreted symbols $\mathcal{EUF}(\Sigma)$ that is well-known to be convex and was shown to have the YMc property in [Yorsh and Musuvathi 2005].

### 5.2. The non-convex case

For *non-convex* theories, the notion of equality interpolation given in this paper is strictly more general than that proposed in the technical report [Yorsh and Musuvathi 2004] accompanying [Yorsh and Musuvathi 2005]. Formally, an arbitrary (either convex or non-convex) theory has the *YM equality interpolating property* (read as the '(general) Yorsh-Musuvathi property') iff it has the quantifier-free interpolation property and satisfies the following condition:

— for every tuples $\underline{x}$, $\underline{z}_1$, $\underline{z}_2$ of variables, further tuples $\underline{y}_1 = y_{11}, \ldots, y_{1n}$, $\underline{y}_2 = y_{21}, \ldots, y_{2n}$ of variables, and pairs $\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1), \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2)$ of conjunctions of literals,

$$\text{if } \delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \bigvee_{i=1}^{n} (y_{1i} = y_{2i}) \text{ holds,}$$

then there exists a tuple $\underline{v}(\underline{x}) = v_1, \ldots, v_n$ of terms such that

$$\delta_1(\underline{x}, \underline{z}_1, \underline{y}_1) \wedge \delta_2(\underline{x}, \underline{z}_2, \underline{y}_2) \vdash_T \bigvee_{i=1}^{n} (y_{1i} = v_i \wedge v_i = y_{2i}).$$

We formally prove here, by using the combination result from [Yorsh and Musuvathi 2004], that the notion of YM equality interpolation *implies* that of equality interpolation of Definition 4.1. As a preliminary observation, notice that (trivially) if a convex theory has the YMc property, then it is also YM equality interpolating. Since $\mathcal{EUF}(\Sigma)$ is convex and has the YMc property (as shown in [Yorsh and Musuvathi 2005]), it is YM equality interpolating. Now if a theory $T$ is YM equality interpolating and if $\Sigma$ is a signature disjoint from that of $T$, then $T \cup \mathcal{EUF}(\Sigma)$ is quantifier-free interpolating since the component theories are signature disjoint and have the YM equality interpolating property and the combination result in [Yorsh and Musuvathi 2004] applies. As a consequence, $T$ is strongly sub-amalgamable (Theorem 3.10) and also equality interpolating in our sense (Theorem 4.2).

However, the converse does not hold as our notion is *strictly weaker* than that of YM equality interpolation. To see this, we exhibit a (non-convex) theory $T_{cex}$ that has the

strong sub-amalgamation property but is not YM equality interpolating. Let the signature of $T_{cex}$ contain three propositional letters $p_1, p_2$ and $p_3$, three constant symbols $c_1, c_2$, and $c_3$, and a unary predicate $Q$. $T_{cex}$ is such that the following properties are its axioms:

— 'exactly one among $p_1, p_2$ and $p_3$ holds',
— '$c_1, c_2$, and $c_3$ are distinct',
— '$Q(x)$ holds for no more than one $x$', and
— '$p_i \rightarrow Q(c_i)$ holds' for $i = 1, 2, 3$.

It is easy to see that $T_{cex}$ is stably infinite and has the strong sub-amalgamation property. $T_{cex}$ is non-convex since $Q(x) \wedge y_1 = c_1 \wedge y_2 = c_2 \wedge y_3 = c_3$ implies the disjunction $x = y_1 \vee x = y_2 \vee x = y_3$ without implying any single disjunct. Now, notice that $Q(x) \wedge Q(y) \vdash_{T_{cex}} x = y$. According to the definition of the YM equality interpolating property (see above), there should be a *single* ground term $v$ such that $Q(x) \wedge Q(y) \vdash_{T_{cex}} x = v \wedge y = v$. This cannot be the case since we must choose among one of the three constants $c_1, c_2, c_3$ to find such a term $v$ and none of these choices fits our purposes. Hence, $T_{cex}$ is not YM equality interpolating although it has the strong sub-amalgamation property and hence it is equality interpolating according to Definition 4.1.

To conclude the comparison with [Yorsh and Musuvathi 2004], we point out that the scope of applicability of our result about the modularity of theories admitting quantifier-free interpolation (Corollary 3.9 above) is *broader* than the one in [Yorsh and Musuvathi 2004]. This is so because, as shown above, our notion of equality interpolation is *strictly weaker* than that of YM equality interpolation in [Yorsh and Musuvathi 2004].

### 5.3. Inadequacy of YMc condition in the non-convex case

By defining a suitable theory, we now show that the the YMc property (see Definition 5.1) is not sufficient to guarantee the modularity of quantifier-free interpolation for non-convex theories. Intuitively, the reason is that disjunctions of equalities must be propagated in the non-convex case and the convex formulation of the equality interpolation property does not consider them.

We define the theory $CL$ of *cuff links* containing at most one pair of *golden* cuff links. Formally, the signature $\Sigma_{CL}$ of $CL$ contains constant $c_0$,[13] a unary function symbol $(-)'$, and a unary predicate symbol $G$. The axioms of $CL$ are the following:

— $\forall x.\ x'' = x, \quad \forall x.\ x \neq x'$ saying that $(-)'$ denotes the 'twin' cuff link,
— $\forall x.\ G(x) \leftrightarrow G(x')$ saying that twin cuff links are either golden or not, and
— last one saying that there is at most one pair of golden cuff links, namely

$$\forall x \forall y.\ G(x) \wedge G(y) \rightarrow x = y \vee x' = y. \tag{23}$$

By using the last axiom, it is easy to realize that $CL$ is not convex.

LEMMA 5.3. *On the one hand, $CL$ has the quantifier-free interpolation property, because it has the sub-amalgamation property. On the other hand, $CL$ does not enjoy the strong sub-amalgamation property.*

PROOF. That the sub-amalgamation property holds is quite clear: suppose we are given models $\mathcal{M}_1, \mathcal{M}_2$ of $CL$ sharing the substructure $\mathcal{A}$ (as a side remark, notice that $\mathcal{A}$ is also a model of $CL$ because $CL$ is universal). As usual, we assume that the intersection of the supports of $\mathcal{M}_1$ and $\mathcal{M}_2$ is the support of $\mathcal{A}$. To amalgamate $\mathcal{M}_1, \mathcal{M}_2$

[13] This is introduced just to make sure that $\Sigma_{CL}$ is non-empty.

over $\mathcal{A}$, it is sufficient to take the union of the supports of $\mathcal{M}_1$ and $\mathcal{M}_2$, with just one proviso: if $\mathcal{M}_1, \mathcal{M}_2$ both contain a pair of golden cuff links that is not from $\mathcal{A}$, then such pairs must be merged (the need of such merging is precisely what shows that strong sub-amalgamation fails). □

PROPOSITION 5.4. *$CL$ has the YMc property.*

PROOF. We shall work with free constants instead of with variables. Consider finite sets of ground literals $A$, $B$ in the signature $\Sigma_{CL}$ enriched with additional free constants (let $\Sigma_A$ be the signature of $A$ and $\Sigma_B$ be the signature of $B$). We call $AB$-common the ground terms built up from free constants occurring both in $A$ and in $B$; ground terms built up from constants occurring in $A$ but not in $B$ are called $A$-strict ($B$-strict ground terms are defined symmetrically). We call a ground term or literal *pure* iff it is either from $\Sigma_A$ or from $\Sigma_B$. We argue by contraposition. Suppose that, for an $A$-strict constant $a$ and a $B$-strict constant $b$, there is no $AB$-common ground term $t$ such that $A \cup B \vdash_{CL} t = a \wedge t = b$; we show that $A \cup B \not\vdash_{CL} a = b$ by exhibiting a $\Sigma_A \cup \Sigma_B$-model $\mathcal{M}$ of $CL$ such that $\mathcal{M} \models A, \mathcal{M} \models B$ and $\mathcal{M} \not\models a = b$.

We can freely make further assumptions on the sets $A, B$. First, we can assume that there is at least one $AB$-common ground term (because $\Sigma_{CL}$ contains at least the constant $c_0$), that terms like $d''$ do not occur in $A \cup B$ (as they can be simplified to $d$), and that if a term occurs in $A \cup B$, so does its twin term (here the twin of a constant $d$ is $d'$ and the twin of $d'$ is $d$).[14] Second, since the number of $\Sigma_A$-ground literals is finite (modulo the identification of a term like $t''$ with $t$), we can assume that *if a $\Sigma_A$-ground literal is entailed (modulo $CL$) by $A \cup B$, then it actually occurs in $A$ (and similarly for $B$)*: the addition of such entailed literals does not in fact compromise our claim, because if the claim holds for the $A, B$ so modified, it holds for the original $A, B$ too. So, let us make the above assumptions. Notice that (since there is at least one ground $AB$-common term), our hypotheses imply that $A \cup B$ is $CL$-consistent, so no pair of contradictory literals can be contained in the union of $A$ and $B$.

We can split the ground terms occurring in $\Sigma_A \cup \Sigma_B$ into equivalence classes (similarly to what happens in congruence closure algorithms), according to the equivalence relation that holds among $d_1$ and $d_2$ iff (i) either they both occur in $A$ and $d_1 = d_2 \in A$, or (ii) they both occur in $B$ and $d_1 = d_2 \in B$, or (iii) $d_1$ is $A$-strict, $d_2$ is $B$-strict and there exists an $AB$-common $t$ such that $d_1 = t \in A$, $d_2 = t \in B$, or (iv) $d_1$ is $B$-strict, $d_2$ is $A$-strict and there exists an $AB$-common $t$ such that $d_1 = t \in B$, $d_2 = t \in A$. Notice that, because of our assumptions, the equivalence class of $a$ is different from the equivalence class of $b$.

Since there are no contradictory literals in $A \cup B$, we can build a $\Sigma_A \cup \Sigma_B$-structure $\mathcal{A}$ in which all literals from $A \cup B$ are true: the support of $\mathcal{A}$ is formed by the above equivalence classes, a free constant is interpreted as the equivalence class it belongs to, the twin $C'$ of an equivalence class $C$ is the equivalence class formed by the twin terms of the terms belonging to $C$; moreover, $C$ is a golden cuff link in $\mathcal{A}$ iff $G(t) \in A \cup B$ (here $t$ is any term belonging to $C$). Notice that $\mathcal{A} \not\models a = b$. However, we are not done, because $\mathcal{A}$ may not be a model of $CL$: the reason is that there might be more than one golden pair of cuff links. We now show how to merge all golden pairs of cuff links of $\mathcal{A}$ and get a model $\mathcal{M}$ of $CL$ having the required properties, namely such that $\mathcal{M} \models A, \mathcal{M} \models B$ and $\mathcal{M} \not\models a = b$.

Consider two different pairs of golden cuff links $C, C'$ and $D, D'$ (when we say that they are different as pairs of cuff links, we mean that $C$ is different from both $D$ and $D'$). We claim that if we merge $C$ with $D$ and $C'$ with $D'$ as equivalence classes (i.e. if we identify them as elements from the support of $\mathcal{A}$), we still have that the literals from

[14] To ensure this, we can just add literals like $d' = d'$ or $d = d$ to $A$ or $B$, if needed.

$A$ and $B$ are true. In fact, this could possibly be not the case if there are $t \in C, u \in D$ such that $t \neq u \in A \cup B$. However, literals in $A \cup B$ are all pure, so that either $t, u \in \Sigma_A$ or $t, u \in \Sigma_B$. Suppose $t, u \in \Sigma_A$ (the other case is symmetric); by the construction of $\mathcal{A}$ and since $C, D$ are golden, we have that $G(t), G(u) \in A \cup B$ and hence (by (23)) the entailed literal $t = u'$ belongs to $A$, so that $C = D'$ which means that $C, C'$ and $D, D'$ are not different pairs of cuff links.

In conclusion, whenever we pick two different pairs of golden cuff links $C, C'$ and $D, D'$ from the support of $\mathcal{A}$, we can merge $C$ with $D$ and $D$ with $D'$, without compromising the truth of $A \cup B$; notice, however, that we can make the symmetric operation and merge $C$ with $D'$ and $C'$ with $D$, again keeping the literals in $A \cup B$ true. In the end, we can merge all golden pairs of cuff links into a single one; if $a$ and $b$ belong to $C$ and $D$, respectively, and if $C, D$ are both golden, we can choose the appropriate merging among the two possible ones, so that in the end we have that $D$ is equal to $C'$, which implies that $a$ and $b$ remains interpreted as different elements in the support of the final model. □

The above results and Theorem 3.10 immediately imply the following result.

COROLLARY 5.5. *$CL$ has the quantifier-free interpolation property and the YMc property, but $CL \cup \mathcal{EUF}$ does not have the quantifier-free interpolation property (if the signature of $\mathcal{EUF}$ has at least a unary predicate symbol).*

A concrete counterexample to the quantifier-free interpolation property for the combined theory $CL \cup \mathcal{EUF}(\Sigma)$ can easily be obtained by considering the following mutually unsatisfiable sets of ground literals

$$A := \{G(a), P(a), P(a')\}, \qquad B := \{G(b), \neg P(b), \neg P(b')\}$$

where $\Sigma$ contains the additional (free) predicate $P$.

### 5.4. Equality interpolation and Beth definability

Now we discuss the relationship between the notions introduced in this paper with standard notions in mathematical logic and universal algebra. This complementary material is included for the sake of completeness and can be skipped by readers interested only in combination of interpolation. Key to do this is the notion of *Beth definability* with the classical result in model theory, called *Beth definability theorem* (see, e.g., [Chang and Keisler 1990]). We will show that, (at least) in the convex case, equality interpolation can be interpreted as a variant 'modulo theory' of the Beth definability property, restricted to a suitable class of formulæ. In the non-convex case, we will also define a 'Beth-like' formulation of equality interpolation. To complete the picture, we use the Beth definability formulation of equality interpolation to briefly discuss the relationships between our results and well-known results about strong amalgamation in the algebraic literature.

We begin by adding a further (equivalent) characterization to (i)—(iii) in Theorem 4.2. We restate an extended version of that result here to simplify the task of the reader.

THEOREM 5.6. *The following conditions are equivalent for a theory $T$ having quantifier-free interpolation:*

(i). *$T$ is strongly sub-amalgamable;*
(ii). *$T$ is equality interpolating;*
(iii). *$T$ satisfies the implication* (6) $\Rightarrow$ (7) *(for every $\delta_1, \delta_2$);*
(iv). *for every quantifier-free formula $\delta(\underline{x}, \underline{z}, \underline{y})$ such that*

$$\delta(\underline{x}, \underline{z}', \underline{y}') \wedge \delta(\underline{x}, \underline{z}'', \underline{y}'') \vdash_T \underline{y}' \cap \underline{y}'' \neq \emptyset \tag{24}$$

*there are terms* $\underline{v}(\underline{x})$ *such that*

$$\delta(\underline{x},\underline{z},\underline{y}) \vdash_T \underline{y} \cap \underline{v} \neq \emptyset. \tag{25}$$

PROOF. We already proved (in the previous formulation of Theorem 4.2) that conditions (i)-(ii)-(iii) are all equivalent to each other.

Assume (iv) and (6). Take $\underline{y} := \underline{y}_1, \underline{y}_2$ and put $\delta(\underline{x},\underline{y}) := \delta_1(\underline{x},\underline{y}_1) \wedge \delta_2(\underline{x},\underline{y}_2)$. Now notice that $\delta(\underline{x},\underline{y}'_1,\underline{y}'_2) \wedge \delta(\underline{x},\underline{y}''_1,\underline{y}''_2)$ is

$$\delta_1(\underline{x},\underline{y}'_1) \wedge \delta_2(\underline{x},\underline{y}'_2) \wedge \delta_1(\underline{x},\underline{y}''_1) \wedge \delta_2(\underline{x},\underline{y}''_2);$$

since by (6) we have

$$\delta_1(\underline{x},\underline{y}'_1) \wedge \delta_2(\underline{x},\underline{y}''_2) \vdash_T \underline{y}'_1 \cap \underline{y}''_2 \neq \emptyset$$

a fortiori we get

$$\delta(\underline{x},\underline{y}'_1,\underline{y}'_2) \wedge \delta(\underline{x},\underline{y}''_1,\underline{y}''_2) \vdash_T \underline{y}'_1\underline{y}'_2 \cap \underline{y}''_1\underline{y}''_2 \neq \emptyset, \tag{26}$$

By (iv), there are terms $\underline{v}(\underline{x})$ such that $\delta(\underline{x},\underline{y}_1,\underline{y}_2) \vdash_T \underline{y}_1\underline{y}_2 \cap \underline{v} \neq \emptyset$, which is the same as (7).

For the vice versa, we suppose that (4) $\Rightarrow$ (5) holds. Consider $\delta(\underline{x},\underline{z},\underline{y})$ such that (24) holds. Then, we can find $\underline{v}(\underline{x})$ such that

$$\delta(\underline{x},\underline{z}',\underline{y}') \wedge \delta(\underline{x},\underline{z}'',\underline{y}'') \vdash_T (\underline{y}' \cap \underline{v} \neq \emptyset) \vee (\underline{y}'' \cap \underline{v} \neq \emptyset) \tag{27}$$

holds. Making the substitutions $\underline{z}' \mapsto \underline{z}, \underline{z}'' \mapsto \underline{z}, \underline{y}' \mapsto \underline{y}, \underline{y}'' \mapsto \underline{y}$, this gives precisely (25). □

Condition (iv) above can be interpreted as a 'generalized Beth property.' This becomes clear when considering the class of convex theories. We first restate Proposition 5.2.

PROPOSITION 5.7. *The following conditions are equivalent for a convex theory* $T$ *having quantifier-free interpolation:*

(i). $T$ *is equality interpolating;*
(ii). $T$ *satisfies the implication* (20) $\Rightarrow$ (21) *(for every conjunctions of literals* $\delta_1, \delta_2$*);*
(iii). *for every pair* $\underline{x},\underline{z}$ *of tuples of variables, for every further variable* $y$ *and for every conjunction of literals* $\delta(\underline{x},\underline{z},y)$ *such that*

$$\delta(\underline{x},\underline{z}',y') \wedge \delta(\underline{x},\underline{z}'',y'') \vdash_T y' = y'' \ ,$$

*there is a term* $v(\underline{x})$ *such that*

$$\delta(\underline{x},\underline{z},y) \vdash_T y = v \ .$$

PROOF. As for the proof of the extended version of Theorem 4.2, we already know that (i) and (ii) are equivalent.

Assume that (iii) holds and consider $\delta_1(\underline{x},\underline{z}_1,y_1), \delta_2(\underline{x},\underline{z}_2,y_2)$ satisfying (20). Take $\delta(\underline{x},\underline{z}_1,\underline{z}_2,y) := \delta_1(\underline{x},\underline{z}_1,y) \wedge \delta_2(\underline{x},\underline{z}_2,y)$. Now $\delta(\underline{x},\underline{z}'_1,\underline{z}'_2,y') \wedge \delta(\underline{x},\underline{z}''_1,\underline{z}''_2,y'')$ is

$$\delta_1(\underline{x},\underline{z}'_1,y') \wedge \delta_2(\underline{x},\underline{z}'_2,y') \wedge \delta_1(\underline{x},\underline{z}''_1,y'') \wedge \delta_2(\underline{x},\underline{z}''_2,y''),$$

hence (considering the first and the fourth conjunct) from (20) we get

$$\delta(\underline{x},\underline{z}'_1,\underline{z}'_2,y') \wedge \delta(\underline{x},\underline{z}''_1,\underline{z}''_2,y'') \vdash_T y' = y''.$$

By (iii), there is a term $v(\underline{x})$ such that

$$\delta_1(\underline{x},\underline{z}_1,y) \wedge \delta_2(\underline{x},\underline{z}_2,y) \vdash_T y = v(\underline{x}). \tag{28}$$

Again by (20), we obtain

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2 \wedge \delta_2(\underline{x}, \underline{z}_2, y_1) \ ;$$

thus (taking into account (28) and renaming $y$ to $y_1$) also

$$\delta_1(\underline{x}, \underline{z}_1, y_1) \wedge \delta_2(\underline{x}, \underline{z}_2, y_2) \vdash_T y_1 = y_2 \wedge y_1 = v(\underline{x})$$

and finally (21).

Vice versa, if (ii) holds and we have $\delta(\underline{x}, \underline{z}', y') \wedge \delta(\underline{x}, \underline{z}'', y'') \vdash_T y' = y''$, we can find $v(\underline{x})$ such that

$$\delta(\underline{x}, \underline{z}', y') \wedge \delta(\underline{x}, \underline{z}'', y'') \vdash_T y' = v \wedge y'' = v \ ;$$

applying the substitution $\underline{z}' \mapsto \underline{z}, \underline{z}'' \mapsto \underline{z}, \underline{y}' \mapsto \underline{y}, \underline{y}'' \mapsto \underline{y}$, this gives our claim $\delta(\underline{x}, \underline{z}, y) \vdash_T y = v$. □

A *primitive* formula is obtained from a conjunction of literals by prefixing to it a string of existential quantifiers. We can reformulate condition (iii) from Proposition 5.7 above as follows:

(iii)'. for every tuple of variables $\underline{x}$, for every further variable $y$ and for every *primitive* formula $\theta(\underline{x}, y)$ such that $\theta(\underline{x}, y') \wedge \theta(\underline{x}, y'') \vdash_T y' = y''$, there is a term $v(\underline{x})$ such that $\theta(\underline{x}, y) \vdash_T y = v$.

This is precisely *Beth definability property* [Chang and Keisler 1990] for primitive formulæ, extended so as to take into account the theory $T$; we will call this variant Beth definability modulo $T$. Hence equality interpolation coincides with this variant of 'primitive Beth definability property' for convex theories.

To conclude the comparison with existing classical notions, we make some observations connecting the above result with the algebraically oriented literature (see [Kiss et al. 1982] for a survey and for pointers to relevant papers). In an appropriate context from universal algebra, strong amalgamability is shown to be equivalent to the conjunction of amalgamability and of regularity of epimorphisms (alternatively: monomorphisms). In the same context, unravelling the definitions and using presentations of algebras as quotient of free ones, it is not difficult to realize that the primitive Beth definability property above is equivalent to regularity of monomorphisms. Thus, *our results perfectly match with the algebraic characterization of strong amalgamability*. Our work, however, is *orthogonal* to algebraic and category-theoretic approaches: such approaches are able in fact to prove characterizations of strong amalgamability that work in abstract sufficiently complete/cocomplete categories, including consequently categories having nothing to do with models of first order theories. On the other hand, existence of minimal categorical structures fails in our context as soon as we go beyond the universal Horn case. Thus, the two approaches are incomparable and this is reflected by the different techniques employed (we mostly rely on diagrams and compactness, whereas the category-theoretic approach mostly exploit universal properties).

## 6. AN INTERPOLATION ALGORITHM FOR COMBINATIONS OF THEORIES

The notion of equality interpolation and Corollary 3.9 allow us to establish quantifier-free interpolation for all those theories obtained by combining a theory axiomatizing a container data structure—such as $\mathcal{EUF}$, $\mathcal{RDS}$, $\mathcal{LST}$, or $\mathcal{AX}_{\mathtt{diff}}$—with relevant fragments of Arithmetics—such as $\mathcal{LAR}$, $\mathcal{IDL}$, $\mathcal{UTVPI}$, or $\mathcal{LAI}$. However, just knowing that quantifier-free interpolants exist is not sufficient for many applications such as software model checking. For these, it is crucial to be able to actually compute interpolants for combinations of theories. In the rest of this section, we present an algo-

rithm for doing this that modularly reuses the available interpolation algorithms for the component theories.

To simplify the technical development, we work with ground formulae over signatures expanded with free constants instead of quantifier-free formulae as done above. We use the letters $A, B, \ldots$ to denote finite sets of ground formulae; the logical reading of a set of formulae is the conjunction of its elements. For a signature $\Sigma$ and a set $A$ of formulae, $\Sigma^A$ denotes the signature $\Sigma$ expanded with the free constants occurring in $A$. Let $A$ and $B$ be two finite sets of ground formulae in the signatures $\Sigma^A$ and $\Sigma^B$, respectively, and $\Sigma^C := \Sigma^A \cap \Sigma^B$. Given a term, a literal, or a formula $\varphi$ we call it:

— $AB$-*common* iff it is defined over $\Sigma^C$;
— $A$-*local* (resp. $B$-*local*) if it is defined over $\Sigma^A$ (resp. $\Sigma^B$);
— $A$-*strict* (resp. $B$-*strict*) iff it is $A$-local (resp. $B$-local) but not $AB$-common;
— $AB$-*mixed* if it contains symbols in both $(\Sigma^A \setminus \Sigma^C)$ and $(\Sigma^B \setminus \Sigma^C)$;
— $AB$-*pure* if it does not contain symbols in both $(\Sigma^A \setminus \Sigma^C)$ and $(\Sigma^B \setminus \Sigma^C)$.

(Sometimes in the literature about interpolation, "$A$-local" and "$B$-local" are used to denote what we call here "$A$-strict" and "$B$-strict").

### 6.1. Interpolating meta-rules

The interpolation method for combination of theories relies on the abstract framework—based on 'meta-rules'—introduced in [Bruttomesso et al. 2011b; 2012] (to which, the interested reader is pointed for more details) and used also in [Bruttomesso et al. 2011a]. A meta-rule applies (bottom-up) to a pair $A, B$ of finite sets of ground formulae[15] producing an equisatisfiable pair of sets of formulae. Each meta-rule comes with a proviso for its applicability and an instruction for the computation of the interpolant. As an example, consider the meta-rule (Define0):

$$\frac{A \cup \{a = t\} \mid B \cup \{a = t\}}{A \mid B}$$

*Proviso*: $t$ is $AB$-common, $a$ is fresh
*Instruction*: $\phi' \equiv \phi(t/a)$.

It is not difficult to see that $A \cup B$ is equisatisfiable to $A \cup B \cup \{a = t\}$ since $a$ is a fresh variable that has been introduced to re-name the $AB$-common term $t$ according to the proviso of (Define0). The instruction attached to (Define0) allows for the computation of the interpolant $\phi'$ by eliminating the fresh constant $a$ from the recursively known interpolant $\phi$.

The idea is to build an *interpolating meta-rules refutation* for a given unsatisfiable $A_0 \cup B_0$. Such a refutation is a finite labeled tree having the following properties: (i) nodes are labeled by pairs of finite sets of ground formulae; (ii) the root is labeled by $A_0, B_0$; (iii) the leaves are labeled by a pair $\tilde{A}, \tilde{B}$ such that $\bot \in \tilde{A} \cup \tilde{B}$; (iv) each non-leaf node is the conclusion of a meta-rule and its successors are the premises of that meta-rule (the complete list of meta-rules is in Table I).

Once an interpolating meta-rules refutation has been built, it is possible to recursively compute the interpolant by using (top-down) the instructions attached to the meta-rules in the tree as stated in the following result.

PROPOSITION 6.1 ([BRUTTOMESSO ET AL. 2011B; 2012]). *If there exists an interpolating meta-rules refutation for $A_0, B_0$ then there is a quantifier-free interpolant for $A_0, B_0$ (i.e., there exists a quantifier-free $AB$-common sentence $\phi$ such that $A_0 \vdash \phi$ and $B_0 \wedge \phi \vdash \bot$). The interpolant $\phi$ is recursively computed by applying the relevant interpolating instructions of the meta-rules in Table I.*

[15] In [Bruttomesso et al. 2011b; 2012; 2011a], meta-rules manipulate pairs of finite sets of literals instead of pairs of finite sets of ground formulae; the difference is immaterial.

Table I. Interpolating Meta-rules: each rule has a proviso $Prov.$ and an instruction $Instr.$ for recursively computing the new interpolant $\phi'$ from the old one(s) $\phi, \phi_1, \ldots, \phi_k$. Meta-rules are applied *bottom-up* and interpolants are computed *top-down*. Meta-rules Disjunction1/2 generates $n$ braching nodes (one for each $k = 1, \ldots, n$).

| Close1 | Close2 | Propagate1 | Propagate2 |
|---|---|---|---|
| $\dfrac{}{A \mid B}$ | $\dfrac{}{A \mid B}$ | $\dfrac{A \mid B \cup \{\psi\}}{A \mid B}$ | $\dfrac{A \cup \{\psi\} \mid B}{A \mid B}$ |
| *Prov.*: $A$ is unsat.<br>*Instr.*: $\phi' \equiv \bot$. | *Prov.*: $B$ is unsat.<br>*Instr.*: $\phi' \equiv \top$. | *Prov.*: $A \vdash \psi$ and $\psi$ is $AB$-common.<br>*Instr.*: $\phi' \equiv \phi \wedge \psi$. | *Prov.*: $B \vdash \psi$ and $\psi$ is $AB$-common.<br>*Instr.*: $\phi' \equiv \psi \rightarrow \phi$. |

| Define0 | Define1 | Define2 |
|---|---|---|
| $\dfrac{A \cup \{a = t\} \mid B \cup \{a = t\}}{A \mid B}$ | $\dfrac{A \cup \{a = t\} \mid B}{A \mid B}$ | $\dfrac{A \mid B \cup \{a = t\}}{A \mid B}$ |
| *Prov.*: $t$ is $AB$-common, $a$ fresh.<br>*Instr.*: $\phi' \equiv \phi(t/a)$. | *Prov.*: $t$ is $A$-local and $a$ is fresh.<br>*Instr.*: $\phi' \equiv \phi$. | *Prov.*: $t$ is $B$-local and $a$ is fresh.<br>*Instr.*: $\phi' \equiv \phi$. |

| Disjunction1 | Disjunction2 |
|---|---|
| $\dfrac{\cdots \quad A \cup \{\psi_k\} \mid B \quad \cdots}{A \mid B}$ | $\dfrac{\cdots \quad A \mid B \cup \{\psi_k\} \quad \cdots}{A \mid B}$ |
| *Prov.*: $\bigvee_{k=1}^{n} \psi_k$ is $A$-local and $A \vdash \bigvee_{k=1}^{n} \psi_k$.<br>*Instr.*: $\phi' \equiv \bigvee_{k=1}^{n} \phi_k$. | *Prov.*: $\bigvee_{k=1}^{n} \psi_k$ is $B$-local and $B \vdash \bigvee_{k=1}^{n} \psi_k$.<br>*Instr.*: $\phi' \equiv \bigwedge_{k=1}^{n} \phi_k$. |

| Redplus1 | Redplus2 | Redminus1 | Redminus2 |
|---|---|---|---|
| $\dfrac{A \cup \{\psi\} \mid B}{A \mid B}$ | $\dfrac{A \mid B \cup \{\psi\}}{A \mid B}$ | $\dfrac{A \mid B}{A \cup \{\psi\} \mid B}$ | $\dfrac{A \mid B}{A \mid B \cup \{\psi\}}$ |
| *Prov.*: $A \vdash \psi$ and $\psi$ is $A$-local.<br>*Instr.*: $\phi' \equiv \phi$. | *Prov.*: $B \vdash \psi$ and $\psi$ is $B$-local.<br>*Instr.*: $\phi' \equiv \phi$. | *Prov.*: $A \vdash \psi$ and $\psi$ is $A$-local.<br>*Instr.*: $\phi' \equiv \phi$. | *Prov.*: $B \vdash \psi$ and $\psi$ is $B$-local.<br>*Instr.*: $\phi' \equiv \phi$. |

| ConstElim1 | ConstElim2 | ConstElim0 |
|---|---|---|
| $\dfrac{A \mid B}{A \cup \{a = t\} \mid B}$ | $\dfrac{A \mid B}{A \mid B \cup \{b = t\}}$ | $\dfrac{A \mid B}{A \cup \{c = t\} \mid B \cup \{c = t\}}$ |
| *Prov.*: $a$ is $A$-strict and does not occur in $A, t$.<br>*Instr.*: $\phi' \equiv \phi$. | *Prov.*: $b$ is $B$-strict and does not occur in $B, t$.<br>*Instr.*: $\phi' \equiv \phi$. | *Prov.*: $c, t$ are $AB$-common, $c$ does not occur in $A, B, t$.<br>*Instr.*: $\phi' \equiv \phi$. |

The idea to build the combination algorithm is the following. We design transformations instructions that can be non-deterministically applied to a pair $A_0, B_0$. Each of the transformation instructions is *justified by meta-rules*, in the sense that it is just a special sequence of applications of meta-rules.

The instructions are such that, whenever they are applied exhaustively to a pair such that $A_0 \cup B_0$ is unsatisfiable, they produce a tree which is an interpolating meta-rules refutation for $A_0, B_0$ from which an interpolant can be extracted according to Proposition 6.1.

### 6.2. A quantifier-free interpolating algorithm

Let $T_i$ be a stably-infinite and equality interpolating theory over the signature $\Sigma_i$ such that the $SMT(T_i)$ problem is decidable and $\Sigma_1 \cap \Sigma_2 = \emptyset$ (for $i = 1, 2$). We assume the

availability of algorithms for $T_1$ and $T_2$ that are able not only to compute quantifier-free interpolants but also the tuples $\underline{v}$ of terms in Definition 4.1 for equality interpolation. Since the $SMT(T_i)$ problem is decidable for $i = 1, 2$, it is always possible to build an equality interpolating algorithm by enumeration; in practice, better algorithms can be designed (see [Yorsh and Musuvathi 2005] for $\mathcal{EUF}$, $\mathcal{LST}$, $\mathcal{LAR}$ and Section 4.1 for the possibility to use quantifier elimination to do this).

Let $\Sigma := \Sigma_1 \cup \Sigma_2$, $T := T_1 \cup T_2$, and $A_0, B_0$ be a $T$-unsatisfiable pair of finite sets of ground formulae over the signature $\Sigma^{A_0 \cup B_0}$. As in the Nelson-Oppen combination method [Nelson and Oppen 1979], there is a pre-processing step in which we purify $A_0$ and $B_0$ so as to eliminate the literals which are neither $\Sigma_1$- nor $\Sigma_2$-literals from them. To do this, it is sufficient to repeatedly apply the technique of "renaming terms by constants" described below. Take a term $t$ (occurring in a literal from $A_0$ or from $B_0$), add the equality $a = t$ for a fresh constant $a$ and replace all the occurrences of $t$ by $a$. The transformation can be justified by the following sequence of meta-rules: Define1, Define2, Redplus1, Redplus2, Redminus1, Redminus2. For example, in the case of the renaming of some term $t$ in $A_0$, the meta-rule Define1 is used to add the explicit definition $a = t$ to $A_0$, the meta-rule Redplus1 to add the formula $\phi(a/t)$ for each $\phi \in A_0$, and the meta-rule Redminus1 to remove from $A_0$ all the formulæ $\phi$ in which $t$ occurs (except $a = t$).

Thanks to the above purification preprocessing, *we assume to manipulate pairs* $A, B$ *of sets of ground formulæ where literals built up of only* $\Sigma_1$*- or of only* $\Sigma_2$*-symbols occur* besides free constants, from now on. This invariant will be maintained during the execution of our algorithm. Given such a pair $A, B$, we denote by $A_1$ and $A_2$ the subsets of $\Sigma_1^A$- and $\Sigma_2^A$-formulae belonging to $A$; the sub-sets $B_1$ and $B_2$ of $B$ are defined similarly. Notice that *it is not the case* that $A \equiv A_1 \cup A_2$ and $B \equiv B_1 \cup B_2$, since quantifier-free formulae can mix $\Sigma_1$- and $\Sigma_2$-symbols even if the literals they are built from do not.

Before presenting our interpolation algorithm for the combination of theories, we recall a technique, called *Term Sharing*, from [Bruttomesso et al. 2011b; 2012]. Suppose that $A$ contains a literal $a = t$, where the term $t$ is $AB$-common and the free constant $a$ is $A$-strict (a symmetric technique applies to $B$ instead of $A$). Then it is possible to "make $a$ $AB$-common" in the following way. First, introduce a fresh $AB$-common constant $c$ with the explicit definition $c = t$ (to be inserted both in $A$ and in $B$, as justified by meta-rule (Define0)); then replace the literal $a = t$ by $a = c$ and replace $a$ by $c$ everywhere else in $A$; finally, delete $a = c$ too. The result is a pair $(A, B)$ where basically nothing has changed but $a$ has been renamed to an $AB$-common constant $c$ (the transformation can be easily justified by a suitable subset of the meta-rules).

An $A$-*relevant atom* is either an atomic formula occurring in $A$ or it is an $A$-local equality between free constants; an $A$-*assignment* is a Boolean assignment $\alpha$ to relevant $A$-atoms; such assignment is required to satisfy $A$, seen as a set of propositional formulæ (relevant $B$-atoms and $B$-assignments are defined similarly). Below, we write $\alpha$ to denote both the assignment $\alpha$ itself and the following set of literals: $\{L \,|\, \alpha(L) = \mathit{true}\}$.

We are now in the position to present the collection of transformations that should be applied non-deterministically and exhaustively to a pair of purified sets of ground formulæ. All the transformations below can be justified by meta-rules; the justification is straightforward and is briefly summarized below (for more details, the interested is pointed to [Bruttomesso et al. 2011b; 2012]). In the following, let $i \in \{1, 2\}$ and $X \in \{A, B\}$.

***Terminate***$_i$. If $A_i \cup B_i$ is $T_i$-unsatisfiable and $\bot \notin A \cup B$, use the interpolation algorithm for $T_i$ to find a ground $AB$-common $\theta$ such that $A_i \vdash_{T_i} \theta$ and $\theta \wedge B_i \vdash_{T_i}$

$\bot$; then add $\theta$ and $\bot$ to $B$. This is justified by the meta-rules (Propagate1) and (Redplus2).

***Decide***$_X$. If there is no $X$-assignment $\alpha$ such that $\alpha \subseteq X$, pick such an assignment $\alpha$ (if there are none, add $\bot$ to $X$); then update $X$ to $X \cup \alpha$. This is justified by the meta-rules (Disjunction1) and (Disjunction2).

***Share***$_i$. Let $\underline{a} = a_1, \ldots, a_n$ be the tuple of the current $A$-strict free constants and $\underline{b} = b_1, \ldots, b_m$ be the tuple of the current $B$-strict free constants. Suppose that $A_i \cup B_i$ is $T_i$-satisfiable, but $A_i \cup B_i \cup \{\underline{a} \cap \underline{b} = \emptyset\}$ is $T_i$-unsatisfiable. Since $T_i$ is equality interpolating, there must exist $AB$-common $\Sigma_i$-ground terms $\underline{v} \equiv v_1, \ldots, v_p$ such that

$$A_i \cup B_i \vdash_{T_i} (\underline{a} \cap \underline{v} \neq \emptyset) \vee (\underline{b} \cap \underline{v} \neq \emptyset).$$

Thus the union of $A_i \cup \{\underline{a} \cap \underline{v} = \emptyset\}$ and of $B_i \cup \{\underline{b} \cap \underline{v} = \emptyset\}$ is not $T_i$-satisfiable and invoking the available interpolation algorithm for $T_i$, we can compute a ground $AB$-common $\Sigma_i$-formula $\theta$ such that $A \vdash_{T_i} \theta \vee \underline{a} \cap \underline{v} \neq \emptyset$ and $\theta \wedge B \vdash_{T_i} \underline{b} \cap \underline{v} \neq \emptyset$. We choose among $n * p + m * p$ alternatives in order to non-deterministically update $A, B$. For the first $n * p$ alternatives, we add some $a_i = v_j$ (for $1 \leq i \leq n$, $1 \leq j \leq p$) to $A$. For the last $m * p$ alternatives, we add $\theta$ to $A$ and some $\{\theta, b_i = v_j\}$ to $B$ (for $1 \leq i \leq m$, $1 \leq j \leq p$). This is justified by meta-rules (Disjunction1), (Propagate1), and (Disjunction2). Term sharing is finally applied to the updated pair in order to decrease the number of the $A$-strict or $B$-strict free constants.

Let $\mathsf{CI}(T_1, T_2)$ be the procedure that, once run on an unsatisfiable pair $A_0, B_0$, first purifies it, then non-deterministically and exhaustively applies the transformation rules above, and finally extracts an interpolant by using the instructions associated to the meta-rules.

THEOREM 6.2. *Let $T_1$ and $T_2$ be two signature disjoint, stably-infinite, and equality interpolating theories having decidable SMT problems. Then, $\mathsf{CI}(T_1, T_2)$ is a quantifier-free interpolation algorithm for the combined theory $T_1 \cup T_2$.*

PROOF. Let $A_0, B_0$ be our input $T_1 \cup T_2$-unsatisfiable pair. By repeatedly applying our transformations **Decide**$_X$, **Share**$_i$ and **Terminate**$_i$ to it, we produce a tree $\tau$ (the pairs labeling the successors of a node are the possible outcomes of our transformations, which are non deterministic). Clearly **Decide**$_X$, **Share**$_i$ and **Terminate**$_i$ are satisfiability-preserving, in the sense that a pair to which they are applied is $T_1 \cup T_2$-satisfiable iff one of the outcomes is. As a consequence, by Lemma 6.3 below, $\bot$ must belongs to all pairs labeling the leaves. Thus, since **Decide**$_X$, **Share**$_i$ and **Terminate**$_i$ can all be justified by meta-rules, our tree $\tau$ is an interpolating meta-rules refutation (and we are done by Proposition 6.1), *provided we show that $\tau$ is finite*. Finiteness of $\tau$ is also needed to prove the termination of our algorithm.

We apply König Lemma and show that all branches of $\tau$ are finite. Notice that the transformation **Decide**$_X$ can be applied many times in a branch: this is because **Share**$_i$ introduces a new ground formula $\theta$ and alters the definition of an $A$-relevant and a $B$-relevant atom (it introduces new $AB$-common constants by Term Sharing). However, **Share**$_i$ can be applied only finitely many times, as it decreases the number of $A$-strict or $B$-strict constants. Once **Share**$_i$ is no more applied, just single applications of **Decide**$_A$, **Decide**$_B$, **Terminate**$_i$ are possible. □

LEMMA 6.3. *If rules **Decide**$_X$, **Share**$_i$ and **Terminate**$_i$ do not apply to a pair $A, B$, then $A \cup B$ is $T_1 \cup T_2$-satisfiable, unless $\bot \in A \cup B$.*

PROOF. Let $\underline{a}, \underline{c}$ the free constants occurring in $A$ and $\underline{b}, \underline{c}$ be the free constants occurring in $B$. If the above rules do not apply and $\bot \notin A \cup B$, then $A_i \cup B_i \cup \{\underline{a} \cap \underline{b} = \emptyset\}$

is $T_i$-satisfiable for $i = 1, 2$; moreover $A$ contains an $A$-assignment $\alpha$ and $B$ contains a $B$-assignment $\beta$. This means that $A_1 \cup A_2$ entails $A$ and $B_1 \cup B_2$ entails $B$, so that it is sufficient to show the $T_1 \cup T_2$-satisfiability of $A_1 \cup A_2 \cup B_1 \cup B_2$ only. The latter follows from Lemma 3.6, because the sets

$$\Theta_i \equiv A_i \cup B_i \cup \{\underline{a} \cap \underline{b} = \emptyset\}$$

satisfy the hypothesis of the Lemma. Pick in fact a pair of constants $d_1, d_2$ from $\underline{a}, \underline{b}, \underline{c}$: if they are both from $\underline{a}, \underline{c}$ or both from $\underline{b}, \underline{c}$, then either $d_1 = d_2$ or $d_1 \neq d_2$ belongs to $\Theta_i$, as $\alpha \cup \beta$ has assigned a truth value to $d_1 = d_2$. If one of them is in $\underline{a}$ and the other is in $\underline{b}$, then $d_1 \neq d_2 \in \Theta_i$ by construction. □

Algorithm $\mathsf{CI}(T_1, T_2)$ paves the way to reuse quantifier-free interpolation algorithms for both conjunctions (see, e.g., [Rybalchenko and Sofronie-Stokkermans 2010]) or arbitrary Boolean combinations of literals (see, e.g., [Cimatti et al. 2008]). In particular, the capability of reusing interpolation algorithms that can efficiently handle the Boolean structure of formulae seems to be key to enlarge the scope of applicability of verification methods based on interpolants [McMillan 2011]. Indeed, one major issue to address in order to make $\mathsf{CI}(T_1, T_2)$ practically usable is to eliminate the non-determinism. We believe this is possible by adapting the Delayed Theory Combination approach [Bozzano et al. 2005].

## 7. DISCUSSION

We have studied the modularity of quantifier-free interpolation. We have shown that strong amalgamability is needed to guarantee quantifier-free interpolation in combinations of theories. We have then introduced a more "operational" but equivalent characterization of strong amalgamability, called equality interpolation. This allows us to design a (non-deterministic) combination algorithm that uses as sub-modules the interpolation algorithms of theories being combined. We have also put in context our notion of equality interpolation with recent work in the fields of verification, automated reasoning, and mathematical logic (with some hints to related notions in universal algebra and category theory). This supports the claim that the proposed notion of equality interpolation is "right one," as it strictly generalizes previous notions and guarantees the modularity of quantifier-free interpolation. Further evidence is provided by the fact that the results of this paper cover several results for the quantifier-free interpolation of combinations of theories that are known from the literature such as

— $\mathcal{EUF}$ and $\mathcal{LST}$ [Yorsh and Musuvathi 2005],
— $\mathcal{EUF}$ and $\mathcal{LAR}$ [McMillan 2005a; Cimatti et al. 2008; Rybalchenko and Sofronie-Stokkermans 2010],
— $\mathcal{EUF}$ and $\mathcal{LIA}$ [Brillout et al. 2011],
— $\mathcal{LST}$ with $\mathcal{LAR}$ [Yorsh and Musuvathi 2005], and
— $\mathcal{AX}_{\mathtt{diff}}$ with $\mathcal{IDL}$ [Bruttomesso et al. 2011a].

Additionally, new (to the best of our knowledge) results about the existence of quantifier-free interpolants can also be derived from the results reported here, e.g.,

— $\mathcal{RDS}$ with $\mathcal{LAR}, \mathcal{IDL}, \mathcal{UTVPI}$, $\mathcal{LIA}$ and $\mathcal{AX}_{\mathtt{diff}}$,
— $\mathcal{LST}$ with $\mathcal{IDL}, \mathcal{UTVPI}$, $\mathcal{LIA}$ and $\mathcal{AX}_{\mathtt{diff}}$, and
— $\mathcal{AX}_{\mathtt{diff}}$ with $\mathcal{LAR}, \mathcal{UTVPI}$, and $\mathcal{LIA}$.

*Related work.* In Section 5, we have extensively discussed the closely related work of [Yorsh and Musuvathi 2005], where the authors illustrate a method to derive in-

terpolants in a Nelson-Oppen combination procedure, provided that the component theories satisfy certain hypotheses.

The work in [Bonacina and Johansson 2011], among other contributions, recasts the method of [Yorsh and Musuvathi 2005] in the context of the $DPLL(T)$ paradigm. The alternative combination method in [Goel et al. 2009] has been designed for its being efficiently incorporated in state-of-the-art SMT solvers but is complete only for convex theories.

An interpolating theorem prover is described in [McMillan 2005a], where a sequent-like calculus is used to derive interpolants from proofs in propositional logic, equality with uninterpreted functions, linear rational arithmetic, and their combinations. The "split" prover in [Jhala and McMillan 2006] applies a sequent calculus for the synthesis of interpolants along the lines of that in [McMillan 2005a] and is tuned for predicate abstraction. The tool can handle combinations of theories involving arrays without extensionality and fragments of Linear Arithmetic. The CSISAT [Beyer et al. 2008] permits the computation of quantifier-free interpolants over a combination of $\mathcal{EUF}$ and $\mathcal{LAR}$ refining the combination method in [Yorsh and Musuvathi 2005]. A version of MATHSAT [Cimatti et al. 2008] features interpolation capabilities for $\mathcal{EUF}$, $\mathcal{LAR}$, $\mathcal{IDL}$, $\mathcal{UTVPI}$, and $\mathcal{EUF}+\mathcal{LAR}$ by extending the Delayed Theory Combination method of [Bozzano et al. 2005].

Methods [Kapur et al. 2006; Kovács and Voronkov 2009; Brillout et al. 2011; McMillan 2011] for the computation of quantified interpolants in the combination of the theory of arrays and Presburger Arithmetic have been proposed. Our work focus on quantifier-free interpolants by identifying suitable variants of the component theories; e.g., $\mathcal{AX}_{\mathtt{diff}}$ instead of $\mathcal{AX}_{\mathrm{ext}}$ and $\mathcal{LIA}$ instead of Presburger Arithmetic. Orthogonal to our approach is the work in [Sofronie-Stokkermans 2006] where interpolation algorithm are developed for extensions of convex theories admitting quantifier-free interpolation.

*Future work.* The approach proposed in this paper allows us to give a uniform and coherent view of many results available in the literature and we hope that it will be the starting point for new developments and implementations. For example, Theorem 6.2 is the key to combine the strength of existing interpolating provers such as those considered above and to widen the scope of applicability of available interpolation algorithms to richer combinations of theories. Another interesting line of future work is to design and implement deterministic versions of the combination algorithm in Section 6.

## REFERENCES

F. Alberti, R. Bruttomesso, S. Ghilardi, S. Ranise, and N. Sharygina. 2012a. Lazy Abstraction with Interpolants for Arrays. In *Proc. of LPAR-18 (LNCS)*. 46–61.

F. Alberti, R. Bruttomesso, S. Ghilardi, S. Ranise, and N. Sharygina. 2012b. SAFARI: SMT-Based Abstraction for Arrays with Interpolants. In *Proc. of CAV (LNCS)*. 679–685.

P. D. Bacsich. 1975. Amalgamation properties and interpolation theorems for equational theories. *Algebra Universalis* 5 (1975), 45–55.

D. Beyer, D. Zufferey, and R. Majumdar. 2008. CSIsat: Interpolation for LA+EUF. In *Proc. of CAV (LNCS)*, Vol. 5123. 304–308.

M. P. Bonacina and M. Johansson. 2011. On interpolation in decision procedures. In *Proc. of TABLEAUX'11 (LNCS)*. 1–16.

M. Bozzano, R. Bruttomesso, A. Cimatti, T. Junttila, P. Van Rossum, S. Ranise, and R. Sebastiani. 2005. Efficient Satisfiability Modulo Theories via Delayed Theory Combination. In *CAV'05 (LNCS)*. 335–349.

A. Brillout, D. Kroening, P. Rümmer, and T. Wahl. 2010. An Interpolating Sequent Calculus for Quantifier-Free Presburger Arithmetic . In *IJCAR (LNCS)*. 384–399.

A. Brillout, D. Kroening, P. Rümmer, and T. Wahl. 2011. Beyond quantifier-free interpolation in extensions of Presburger arithmetic. In *Proc. of VMCAI (LNCS)*. 88–102.

R. Bruttomesso, S. Ghilardi, and S. Ranise. 2011a. A Combination of Rewriting and Constraint Solving for the Quantifier-free Interpolation of Arrays with Integer Difference Constraints. In *Proc. of FroCoS (LNCS)*. 103–118.

R. Bruttomesso, S. Ghilardi, and S. Ranise. 2011b. Rewriting-based Quantifier-free Interpolation for a Theory of Arrays. In *Proc. of RTA (Leibniz Int. Proc. in Informatics)*, Vol. 10. Dagstuhl Publishing, 171–186.

R. Bruttomesso, S. Ghilardi, and S. Ranise. 2012. Quantifier-free Interpolation for a Theory of Arrays. *Logical Methods in computer Science* 8, 2 (2012).

C. Chang and J. H. Keisler. 1990. *Model Theory* (third ed.). North-Holland, Amsterdam-London.

A. Cimatti, A. Griggio, and R. Sebastiani. 2008. Efficient Interpolant Generation in Satisfiability Modulo Theories. In *TACAS (LNCS)*. 397–412.

W. Craig. 1957. Three uses of the Herbrand-Gentzen theorem in relating model theory and proof theory. *J. Symb. Log.* (1957), 269–285.

George B. Dantzig and B. Curtis Eaves. 1973. Fourier-Motzkin elimination and its dual. *Journal of Combinatorial Theory* 14 (1973), 288–297.

L. Van Der Dries. *Mathematical Logic Lecture Notes*. Technical Report. Available at http://www.math.uiuc.edu/~vddries/.

Herbert B. Enderton. 1972. *A Mathematical Introduction to Logic*. Academic Press, New York-London.

A. Fuchs, A. Goel, J. Grundy, S. Krstić, and C. Tinelli. 2009. Ground Interpolation for the Theory of Equality. In *TACAS (LNCS)*. 413–427.

S. Ghilardi. 2004. Model Theoretic Methods in Combined Constraint Satisfiability. *Journal of Automated Reasoning* 33, 3-4 (2004), 221–249.

A. Goel, S. Krstić, and C. Tinelli. 2009. Ground Interpolation for Combined Theories. In *Proc. of CADE 22 (LNCS)*. 183–198.

T. Henzinger and K. L. McMillan R. Jhala, R. Majumdar. 2004. Abstractions from Proofs. In *POPL*. 232–244.

R. Jhala and K. L. McMillan. 2006. A Practical and Complete Approach to Predicate Refinement. In *Proc. of TACAS (LNCS)*. 459–473.

Bjarni Jónsson. 1956. Universal relational systems. *Math. Scand.* 4 (1956), 193–208.

D. Kapur, R. Majumdar, and C. Zarba. 2006. Interpolation for Data Structures. In *SIGSOFT'06/FSE-14*. 105–116.

E. W. Kiss, L. Márki, P. Pröhle, and W. Tholen. 1982. Categorical algebraic properties. A compendium on amalgamation, congruence extension, epimorphisms, residual smallness, and injectivity. *Studia Sci. Math. Hungar.* 18, 1 (1982), 79–140.

L. Kovács and A. Voronkov. 2009. Finding Loop Invariants for Programs over Arrays Using a Theorem Prover. In *FASE (LNCS)*. 470–485.

L. Kovcs and A. Voronkov. 2009. Interpolation and Symbol Elimination. In *Proc. of CADE-22 (LNCS)*. 199–213.

A. I. Mal'cev. 1962. Axiomatizable Classes of Locally Free Algebras of Certain Types. *Sibirsk. Mat. Ž.* 3 (1962), 729–743.

K. L. McMillan. 2004. An Interpolating Theorem Prover. In *Proc. of TACAS (LNCS)*. 16–30.

K. L. McMillan. 2005a. An Interpolating Theorem Prover. *Theor. Comput. Sci.* 345, 1 (2005), 101–121. DOI: http://dx.doi.org/10.1016/j.tcs.2005.07.003

K. L. McMillan. 2005b. Applications of Craig Interpolation to Model Checking. In *Proc. of TACAS (LNCS)*. 1–12.

K. L. McMillan. 2011. Interpolants from Z3 proofs. In *Proc. of FMCAD*. 19–27.

G. Nelson and D. C. Oppen. 1979. Simplification by Cooperating Decision Procedures. *ACM Transactions on Programming Languages and Systems* 1, 2 (1979), 245–57.

D. C. Oppen. 1980. Reasoning about Recursively Defined Data Structures. *J. ACM* 27 (1980), 403–411.

Claus Michael Ringel. 1972. The intersection property of amalgamations. *J. Pure Appl. Algebra* 2 (1972), 341–342.

A. Rybalchenko and V. Sofronie-Stokkermans. 2010. Constraint Solving for Interpolation. *J. of Symbolic Computation* 45, 11 (2010), 1212–1233.

V. Sofronie-Stokkermans. 2006. Interpolation in Local Theory Extensions. In *Proc. of IJCAR'06 (LNCS)*, Vol. 4130. 235–250.

C. Tinelli and M. T. Harandi. 1996. A new correctness proof of the Nelson-Oppen combination procedure. In *Proc. of FroCoS 1996*. 103–119.

G. Yorsh and M. Musuvathi. 2004. *A combination method for generating interpolants*. Technical Report. Microsoft Research. Technical Report MSR-TR-2004-108.

G. Yorsh and M. Musuvathi. 2005. A combination method for generating interpolants. In *Proc. of CADE-20*. 353–368.

Received December 2012; revised May 2013; accepted June 2013