VARIANTS OF VARIANTS AND THE FINITE VARIANT PROPERTY

ANDREW CHOLEWA, JOSÉ MESEGUER, AND SANTIAGO ESCOBAR

Abstract. Variants and the finite variant property were originally introduced about a decade ago by Hurbert Comon-Lundh and Stéphanie Delaune to reason about equational theories that commonly appear in cryptographic protocol analysis. Since that time, two additional notions of variants have been developed: one by Santiago Escobar, José Meseguer, and Ralf Sasse, and one by Ştefan Ciobâcă. Though it seems intuitively clear that all three notions capture the same essential idea, their relationships to each other have never been rigorously analyzed. Therefore, we decided to do just that. In the process, we encountered an unexpected subtlety with respect to the finite variant property, the term signature of a theory, and the boundedness property. We also provide a simple semi-decision procedure for checking if an equational theory has the finite variant property, by analzying terms of the form $f(x_1 : s_1, \ldots, x_n : s_n)$ for all function symbols $f : s_1, \ldots, s_n \to s$ in the signature, where $x_1, \ldots, x_n$ are distinct variables.

1. Introduction

Just under a decade ago, Hubert Comon-Lundh and Stéphanie Delaune introduced the notion of a variant of a term t with respect to an equational theory in [2]. A variant is a normalized term, t', and a normalized substitution, θ, such that $t\theta$ rewrites to t'. In other words, variants are patterns describing the canonical forms of instances of a term. Furthermore, for some theories we can guarantee that every term t will have a finite number of most general variants. This idea is known as the Finite Variant Property (FVP). Variants have proven to be quite useful for symbolic reasoning about rewrite theories. In particular, when a theory has FVP, an infinite number of rewrite sequences can be reasoned about using only a finite number of terms and substitutions.

The finite variant property is useful enough that plenty of work has been dedicated to determining if a given equational theory has FVP, e.g. in [2] and [4]. This report builds on these earlier results by introducing a simple semi-decision procedure for checking the finite variant property. The correctness of this procedure depends on the equivalence of boundedness and the finite variant property proved in [2]. Boundedness essentially says that for a term t, there is a bound n such that for any rewrite sequence starting at $t\sigma$, for σ a normalized substitution, a canonical form can be reached in at most n steps.

Due to the power of the finite variant property, variants have been applied in several different situations. As a result, several different notions of variants have arisen. There are three notions in particular that we are aware of. First is the original definition proposed by Comon-Lundh and Delaune, in which a variant is

Date: February 11, 2014.
Partially supported by NSF Grants CCF 09-05584, and CNS 13-19109.

characterized as the normalized term t'. Second is a version first proposed by Santiago Escobar, José Meseguer, and Ralf Sasse in [4], and expanded upon in [6]. Escobar et al. view variants as pairs (t', θ) of the normalized term t' and substitution θ such that $t\theta$ rewrites to t'. Third is a version discussed by Ştefan Ciobâcă in his PhD thesis, that characterizes a variant in terms of the substitution θ [3].

It seems intuitively obvious that these various definitions of variants are equivalent. However, a careful reading of the chapter in Ciobâcă's thesis that explores variants throws some doubt on this. First, he introduces the notion of a strongly complete set of variants rather than a complete set of variants. Furthermore, he argues that a strongly complete set of variants is strictly stronger than a complete set of variants. However, the property that distinguishes a strongly complete set of variants from a complete set of variants is also a part of the definition of a complete set of variants used by Escobar et al. in [4]. This would seem to suggest that perhaps the various definitions are not actually equivalent. Even more distressing, the example used by Ciobâcă to argue that strong completeness is stronger than completeness appears to contradict the equivalence of boundedness and the finite variant property.

Considering the importance of the notion of variants, this work investigates the relationship between these different definitions in greater detail. We show that when the signature of an equational theory is fixed, then the notion of variants developed by Comon-Lundh and Delaune and the notion used by Ciobâcă are equivalent, while the notion developed by Escobar, Meseguer, and Sasse is strictly stronger. Furthermore, the strong variant property proposed by Ciobâcă is shown to be equivalent to the notion of the finite variant property used by Escobar et al. If however, the signature of an equational theory is allowed to be expanded (in particular, if one can add a free binary function symbol) then all these notions of variants, and their associated finite variant properties, are equivalent. Meanwhile, the finite variant property as defined by Comon-Lundh and Delaune is not equivalent to boundedness when the signature of a theory is fixed. However, when the signature can be expanded, then the equivalence of the finite variant property and boundedness holds.

Note that the treatements of variants by Comon-Lundh and Delaune, and Ciobâcă are untyped, whereas the treatment by Escobar, Meseguer, and Sasse is order-sorted. This may seem to be just a manner of taste, but it is not. It has become increasingly clear in many examples that an order-sorted typed structure can greatly increase the chances of a theory having the FVP. Since the order-sorted setting is more general and contains the untyped case as the very special case in which there is only one sort and no subsorts, we use throughout the more general and useful order-sorted setting.

2. Preliminaries

We follow the classical notation and terminology from [13] for term rewriting, and from [10] for rewriting logic and order-sorted notions. We assume an order-sorted signature $\mathbf{\Sigma} = (\mathsf{S}, \leq, \Sigma)$ with poset of sorts $(\mathsf{S}, \leq)$ and such that for each sort $\mathsf{s} \in \mathsf{S}$ the connected component of s in $(\mathsf{S}, \leq)$ has a top sort, denoted $[\mathsf{s}]$, and all $f : \mathsf{s_1} \cdots \mathsf{s_n} \to \mathsf{s}$ with $n \geq 1$ have a top sort overloading $f : [\mathsf{s_1}] \cdots [\mathsf{s_n}] \to [\mathsf{s}]$. We also assume an S-sorted family $\mathcal{X} = \{\mathcal{X}_\mathsf{s}\}_{\mathsf{s} \in \mathsf{S}}$ of disjoint variable sets with each $\mathcal{X}_\mathsf{s}$ countably infinite. $\mathcal{T}_\Sigma(\mathcal{X})_\mathsf{s}$ is the set of terms of sort s, and $\mathcal{T}_{\Sigma,\mathsf{s}}$ is the set of ground terms of sort s. We write $\mathcal{T}_{\Sigma}(\mathcal{X})$ and $\mathcal{T}_{\Sigma}$ for the corresponding order-sorted term algebras. For a term t, $\mathit{Var}(t)$ denotes the set of all variables in t.

Positions are represented by sequences of natural numbers denoting an access path in the term when viewed as a tree. The top or root position is denoted by the empty sequence Λ. We define the relation $p \leq q$ between positions as $p \leq p$ for any p; and $p \leq p.q$ for any p and q. Given $U \subseteq \Sigma \cup \mathcal{X}$, $\mathit{Pos}_U(t)$ denotes the set of positions of a term t that are rooted by symbols or variables in U. The set of positions of a term t is written $\mathit{Pos}(t)$, and the set of non-variable positions $\mathit{Pos}_{\Sigma}(t)$. The subterm of t at position p is $t|_p$ and $t[u]_p$ is the term t where $t|_p$ is replaced by u.

A *substitution* $\sigma \in \mathcal{S}\mathit{ubst}(\Sigma, \mathcal{X})$ is a sorted mapping from a finite subset of $\mathcal{X}$ to $\mathcal{T}_{\Sigma}(\mathcal{X})$. Substitutions are written as $\sigma = \{X_1 \mapsto t_1, \ldots, X_n \mapsto t_n\}$ where the domain of σ is $\mathit{Dom}(\sigma) = \{X_1, \ldots, X_n\}$ and the set of variables introduced by terms $t_1, \ldots, t_n$ is written $\mathit{Ran}(\sigma)$. The identity substitution is *id*. Substitutions are homomorphically extended to $\mathcal{T}_{\Sigma}(\mathcal{X})$. The application of a substitution σ to a term t is denoted by $t\sigma$. For simplicity, we assume that every substitution is idempotent, i.e., σ satisfies $\mathit{Dom}(\sigma) \cap \mathit{Ran}(\sigma) = \emptyset$. Substitution idempotency ensures $t\sigma = (t\sigma)\sigma$. The restriction of σ to a set of variables V is $\sigma|_V$; sometimes we write $\sigma|_{t_1,\ldots,t_n}$ to denote $\sigma|_V$ where $V = \mathit{Var}(t_1) \cup \cdots \cup \mathit{Var}(t_n)$. Composition of two substitutions is denoted by $\sigma\sigma'$. Combination of two substitutions is denoted by $\sigma \cup \sigma'$. We call an idempotent substitution σ a variable *renaming* if there is another idempotent substitution σ^{-1} such that $(\sigma\sigma^{-1})|_{\mathit{Dom}(\sigma)} = \mathit{id}$.

A Σ-*equation* is an unoriented pair $t = t'$, where $t, t' \in \mathcal{T}_{\Sigma}(\mathcal{X})_{\mathsf{s}}$ for some sort $\mathsf{s} \in \mathsf{S}$. Given Σ and a set $\mathcal{E}$ of Σ-equations, order-sorted equational logic induces a congruence relation $=_{\mathcal{E}}$ on terms $t, t' \in \mathcal{T}_{\Sigma}(\mathcal{X})$ (see [11]). Throughout this paper we assume that $\mathcal{T}_{\Sigma,\mathsf{s}} \neq \emptyset$ for every sort s, because this affords a simpler deduction system. An *equational theory* $(\Sigma, \mathcal{E})$ is a pair with Σ an order-sorted signature and $\mathcal{E}$ a set of Σ-equations.

The $\mathcal{E}$-*subsumption* preorder $\sqsubseteq_{\mathcal{E}}$ (or just $\sqsubseteq$ if $\mathcal{E}$ is understood) holds between $t, t' \in \mathcal{T}_{\Sigma}(\mathcal{X})$, denoted $t \sqsubseteq_{\mathcal{E}} t'$ (meaning that t' is *more general* than t modulo $\mathcal{E}$), if there is a substitution σ such that $t =_{\mathcal{E}} t'\sigma$; such a substitution σ is said to be an $\mathcal{E}$-*match* from t to t'. Relation $\sqsubseteq_{\mathcal{E}}$ is extended to substitutions in a similar way. For substitutions σ, ρ and a set of variables V we define $\sigma|_V =_{\mathcal{E}} \rho|_V$ if $x\sigma =_{\mathcal{E}} x\rho$ for all $x \in V$; $\sigma|_V \sqsubseteq_{\mathcal{E}} \rho|_V$ if there is a substitution η such that $\sigma|_V =_{\mathcal{E}} (\rho\eta)|_V$.

A *rewrite rule* is an oriented pair $l \to r$, where $\mathit{Var}(r) \subseteq \mathit{Var}(l)$ and $l, r \in \mathcal{T}_{\Sigma}(\mathcal{X})_{\mathsf{s}}$ for some sort $\mathsf{s} \in \mathsf{S}$. An *(unconditional) order-sorted rewrite theory* is a triple (Σ, B, R) with Σ an order-sorted signature, B a set of Σ-equations, and R a set of rewrite rules. The rewriting relation on $\mathcal{T}_{\Sigma}(\mathcal{X})$, written $t \to_R t'$ or $t \to_{p,R} t'$ holds between t and t' iff there exist $p \in \mathit{Pos}_{\Sigma}(t)$, $l \to r \in R$ and a substitution σ, such that $t|_p = l\sigma$, and $t' = t[r\sigma]_p$. The subterm $t|_p$ is called a *redex*. The relation $\to_{R/B}$ on $\mathcal{T}_{\Sigma}(\mathcal{X})$ is $=_B;\to_R;=_B$. Note that $\to_{R/B}$ on $\mathcal{T}_{\Sigma}(\mathcal{X})$ induces a relation $\to_{R/B}$ on the free (Σ, B)-algebra $\mathcal{T}_{\Sigma/B}(\mathcal{X})$ by $[t]_B \to_{R/Ax} [t']_{Ax}$ iff $t \to_{R/Ax} t'$. The transitive and reflexive closure of $\to_{R/B}$ is denoted $\to^*_{R/B}$. We say that a term t is $\to_{R/B}$-irreducible (or just R/B-irreducible) if there is no term t' such that $t \to_{R/B} t'$.

For a rewrite rule $l \to r$, we say that it is *sort-decreasing* if for each substitution σ, we have $r\sigma \in \mathcal{T}_{\Sigma}(\mathcal{X})_{\mathsf{s}}$ implies $l\sigma \in \mathcal{T}_{\Sigma}(\mathcal{X})_{\mathsf{s}}$. We say a rewrite theory (Σ, B, R) is sort-decreasing if all rules in R are. For a Σ-equation $t = t'$, we say that it is *regular* if $Var(t) = Var(t')$, and it is *sort-preserving* if for each substitution σ, we have $t\sigma \in \mathcal{T}_\Sigma(\mathcal{X})_\mathsf{s}$ implies $t'\sigma \in \mathcal{T}_\Sigma(\mathcal{X})_\mathsf{s}$ and vice versa. We say an equational theory $(\Sigma, \mathcal{E})$ is regular or sort-preserving if all equations in $\mathcal{E}$ are.

For substitutions σ, ρ and a set of variables V we define $\sigma|_V \rightarrow_{R/B} \rho|_V$ if there is $x \in V$ such that $x\sigma \rightarrow_{R/B} x\rho$ and for all other $y \in V$ we have $y\sigma =_B y\rho$. A substitution σ is called *R/B-normalized* (or normalized) if $x\sigma$ is R/B-irreducible for all $x \in V$.

We say that the relation $\rightarrow_{R/B}$ is *terminating* if there is no infinite sequence $t_1 \rightarrow_{R/B} t_2 \rightarrow_{R/B} \cdots t_n \rightarrow_{R/B} t_{n+1} \cdots$. We say that the relation $\rightarrow_{R/B}$ is *confluent* if whenever $t \rightarrow^*_{R/B} t'$ and $t \rightarrow^*_{R/B} t''$, there exists a term t''' such that $t' \rightarrow^*_{R/B} t'''$ and $t'' \rightarrow^*_{R/B} t'''$.

2.1. R, B**-rewriting.** Since B-congruence classes can be infinite, $\rightarrow_{R/B}$-reducibility is undecidable in general. Therefore, R/B-rewriting is usually implemented [8] by R, B-rewriting. We assume the following properties on R and B:

(1) B is regular and sort-preserving; furthermore, for each equation $t = t'$ in B, all variables in $Var(t)$ have a top sort.
(2) B has a finitary and complete unification algorithm.
(3) The rewrite rules R are sort-decreasing, confluent, and terminating.

Definition 1 (Rewriting modulo)**.** [14] Let (Σ, B, R) be an order-sorted rewrite theory satisfying properties (1)–(3). We define the relation $\rightarrow_{R,B}$ on $\mathcal{T}_\Sigma(\mathcal{X})$ by $t \rightarrow_{p,R,B} t'$ (or just $t \rightarrow_{R,B} t'$) iff there is a non-variable position $p \in Pos_\Sigma(t)$, a rule $l \rightarrow r$ in R, and a substitution σ such that $t|_p =_B l\sigma$ and $t' = t[r\sigma]_p$.

Note that, since B-matching is decidable, $\rightarrow_{R,B}$ is decidable. Notions such as confluence, termination, irreducible terms, and normalized substitution, are defined in a straightforward manner for $\rightarrow_{R,B}$. Note that since R is sort-decreasing, confluent, and terminating, i.e., the relation $\rightarrow_{R/B}$ is confluent and terminating, and $\rightarrow_{R,B} \subseteq \rightarrow_{R/B}$, the relation $\rightarrow^!_{R,B}$ is decidable, i.e., it terminates and produces a unique term (up to B-equivalence) for each initial term t, denoted by $t{\downarrow}_{R,B}$. Of course $t \rightarrow_{R,B} t'$ implies $t \rightarrow_{R/B} t'$, but the converse does not need to hold in general. To prove completeness of $\rightarrow_{R,B}$ w.r.t. $\rightarrow_{R/B}$ we need the following additional *coherence* assumption; we refer the reader to [7, 14, 9] for coherence completion algorithms.

(4) $\rightarrow_{R,B}$ is *B-coherent* [8], i.e., $\forall t_1, t_2, t_3$ we have $t_1 \rightarrow_{R,B} t_2$ and $t_1 =_B t_3$ implies $\exists t_4, t_5$ such that $t_2 \rightarrow^*_{R,B} t_4$, $t_3 \rightarrow^+_{R,B} t_5$, and $t_4 =_B t_5$. See Figure 1 for a graphical illustration.

The following theorem in [8, Proposition 1] that generalizes ideas in [12] and has an easy extension to order-sorted theories, links $\rightarrow_{R/B}$ with $\rightarrow_{R,B}$.

Theorem 1 (Correspondence)**.** [12, 8] *Let (Σ, B, R) be an order-sorted rewrite theory satisfying properties* (1)*–*(4)*. Then $t_1 \rightarrow^!_{R/B} t_2$ iff $t_1 \rightarrow^!_{R,B} t_3$, where $t_2 =_B t_3$.*

We say that $\rightarrow_{R,B}$ is *convergent* if it is confluent, terminating, sort decreasing and B-coherent. An order-sorted rewrite theory (Σ, B, R) is convergent (resp. terminating, confluent) if the relation $\rightarrow_{R,B}$ is convergent (resp. terminating, confluent). In a convergent order-sorted rewrite theory, for each term $t \in \mathcal{T}_\Sigma(\mathcal{X})$, there is a unique (up to B-equivalence) R, B-irreducible term t' obtained from t by

$$\begin{tikzcd}
t_1 \arrow[r, "{R,B}"'] \arrow[d, Leftrightarrow, "B"'] & t_2 \arrow[dd, dotted, "{*}_{R,B}"'] \\
t_3 \arrow[d, dotted, "{+}_{R,B}"'] & \\
t_5 \arrow[r, dotted, leftrightarrow, "B"'] & t_4
\end{tikzcd}$$

Figure 1. B-coherence

rewriting to canonical form, which is denoted by $t \rightarrow^{!}_{R,B} t'$, or $t{\downarrow}_{R,B}$ when t' is not relevant. Finally, we provide the notion of decomposition of an equational theory into rules and axioms.

Definition 2 (Decomposition)**.** [5] Let $(\Sigma, \mathcal{E})$ be an order-sorted equational theory. We call (Σ, B, R) a *decomposition* of $(\Sigma, \mathcal{E})$ if $\mathcal{E} = R \cup B$ and (Σ, B, R), with the equation $t = t' \in R$ oriented as rules $t \rightarrow t'$ in (Σ, B, R), is a convergent order-sorted rewrite theory.

Note that we abuse notation and call (Σ, B, R) a decomposition of an order-sorted equational theory $(\Sigma, \mathcal{E})$ even when, strictly speaking $\mathcal{E} \neq R \cup B$ but R is the explicitly extended B-coherent version of a set R' such that $\mathcal{E} = R' \cup B$.

3. Variant Definitions

Variants were originally introduced by Comon-Lundh and Delaune in [2]. However in [2], variants are not defined in terms of a terminating rewriting relation, but rather in terms of an arbitrary well-founded ordering $\geq$ on terms, which is total on ground terms. Furthermore, given a set of equations $\mathcal{E}$, they define $\min_{\geq}(t)$ ($\min(t)$ when $\geq$ is understood) as the smallest term in the $\mathcal{E}$-equivalence of t. This is a generalization of the normal form of a term t with respect to a rewriting relation.

Note that in [2] both the notion of a variant, and a complete set of variants were defined together. We have split them into two definitions for ease of reference.

Definition 3 (Comon-Lundh, Delaune Variants (C-LD-Variants))**.** [2] Let $(\Sigma, \mathcal{E})$ be an order-sorted equational theory with a single sort, and let t and t' be Σ-terms. Then, t' is a Comon-Lundh, Delaune variant (C-LD-variant) of t iff there is a substitution θ such that $t\theta =_{\mathcal{E}} t'$.

Definition 4 (Complete Set of C-LD-Variants)**.** [2] Let $(\Sigma, \mathcal{E})$, and $(\Sigma, \mathcal{E}')$ be equational theories. Let t be a Σ-term. Let S be a subset of the set of all C-LD-variants of t with respect to $\geq$. Then, S is a complete set of C-LD-variants of t modulo $\mathcal{E}'$ iff for every substitution π, there is a term $t' \in S$ and a substitution τ, such that $\min_{\geq}(t\pi) =_{\mathcal{E}'} t'\tau$.

Ciobâcă uses the ideas introduced by Comon-Lundh and Delaune in his PhD thesis, which investigates the security of cryptographic protocols [3]. In his thesis, Ciobaca restricts himself to the case of theories with no axioms. Furthermore, whereas Comon-Lundh and Delaune focus on the term t' that is E-equal to an instance of t, Ciobaca focuses on the substitution θ that instantiates t.

Definition 5 (Ciobâcă Variants (Ciob-Variants))**.** [3] Let $(\Sigma, \emptyset, R)$ be a convergent order-sorted term rewriting system, where Σ has a single sort. Let t be a Σ-term. Then, a Ciobaca variant (Ciob-variant) of t with respect to $(\Sigma, \emptyset, R)$ is a substitution θ, where $t\theta \downarrow R$ is a C-LD-variant.

Definition 6 (Complete Set of Ciobâcă Variants)**.** [3] Let $(\Sigma, \emptyset, R)$ be a convergent term rewriting system, and let t be a Σ-term. Then, a complete set of Ciob-variants of t is a set of substitutions Θ such that for any substitution π, there exists $\theta \in \Theta$ and a substitution τ such that $(t\pi)\downarrow_R = ((t\theta)\downarrow_R)\tau$

Ciobâcă also introduces the notion of a strongly complete set of variants.

Definition 7 (Strongly Complete Set of Ciobâcă Variants)**.** [3] Let $(\Sigma, \emptyset, R)$ be a convergent rewrite system, and let t be a Σ-term. A set of substitutions Θ is called a *strongly complete set of variants* of t iff $Dom(\theta) \subseteq var(t)$ for all $\theta \in \Theta$, Θ is a complete set of Ciob-variants, and for every substitution π, the corresponding substitutions θ and τ have the property that $(\pi|_{var(t)})\downarrow_R = ((\sigma\downarrow_R)\tau)|_{var(t)}$.

Note the additional restriction that $\pi\downarrow_R$ decomposes into $\theta_i\downarrow_R$ and τ.

Finally, we have the notion of variants developed by Escobar, Meseguer, and Sasse in [4]. Rather than focusing exclusively on a term or on a substitution, this definition considers both.

Definition 8 (Escobar, Meseguer, Sasse Variant(EMS-Variant))**.** [4] Given a term t and an order-sorted equational theory $(\Sigma, \mathcal{E})$, (t', θ) is an Escobar Meseguer, Sasse (EMS)-variant of t iff $t\theta =_{\mathcal{E}} t'$, with $Dom(\theta) \subseteq Var(t)$ and $Ran(\theta) \cap Var(t) = \emptyset$.

Definition 9 (Complete Set of EMS-Variants)**.** [4] Let (Σ, B, R) be a decomposition of an order-sorted equational theory $(\Sigma, \mathcal{E})$. A complete set of EMS-variants (up to renaming) of a term t is a subset V of the set of EMS-variants of t such that for every substitution π, there is an EMS-variant $(t', \theta) \in V$ and a substitution τ such that:

(1) t' is R, B irreducible
(2) $(t\theta)\downarrow_{R,B} =_B t'\tau$
(3) $(\pi\downarrow_{R,B})|_{var(t)} =_B (\theta\tau)|_{var(t)}$

Note that condition 3 is the same condition imposed by Ciobâcă's notion of a strongly complete set of variants in Definition 7.

Finally, we have the finite variant property.

Definition 10 (Finite Variant Property)**.** Let (Σ, B, R) be an equational decomposition of equational theory $(\Sigma, \mathcal{E})$. Then, $(\Sigma, \mathcal{E})$ is said to have the finite variant property under (Σ, B, R) in the sense of C-LD (resp. EMS, resp. Ciob) iff for every Σ-term t, there exists a finite complete set of C-LD (resp. EMS, resp. Ciob) variants of t.

We say that $(\Sigma, \mathcal{E})$ has the strong finite variant property in the sense of Ciobâcă (Ciob-SFVP) under (Σ, B, R) iff $B = \emptyset$, and for every Σ-term t there exists a finite strongly complete set of variants of t in the sense of Definition 7.

4. Comon-Lundh, Delaune and Ciobâcă Variants

In this section, we show that when the $\geq$ relation in Definition 3 is a convergent rewrite relation, then the C-LD and Ciob-variants are equivalent in the sense that a theory has the C-LD finite variant property iff it also has the Ciob finite variant property.

Theorem 2 (Equivalence of C-LD and Ciob-Variants)**.** *Let $(\Sigma, \emptyset, R)$ be an equational decomposition of the equational theory $(\Sigma, \mathcal{E})$, where Σ contains a single sort. Let t be a Σ-term, and X be a set of variables with $var(t) \subseteq X$. Let $\geq$ be $\rightarrow_R$, and thus $\min_{\geq}(t) = t\downarrow_R$. Then, there exists a finite complete set of Ciob-variants of t iff there exists a finite complete set of C-LD-variants of t.*

Proof. Let C be a finite complete set of Ciob-variants of t. Let $L = \{t\sigma \downarrow_R | \sigma \in C\}$. Clearly, since C is finite, and $t\sigma \downarrow_R$ is unique, L must be finite. Furthermore, observe that since $(\Sigma, \emptyset, R)$ is an equational decomposition of $(\Sigma, \mathcal{E})$, $t\sigma\downarrow_R =_{\mathcal{E}} t\sigma$. So, each $t\sigma\downarrow$ is a C-LD variant of t.

Now, let π be a substitution. Since C is complete, there exists $\sigma \in C$, and a substitution τ such that $t\pi \downarrow_R = (t\sigma) \downarrow_R \tau$. So L is complete.

Let L be a finite complete set of C-LD-variants. Let $C_{t'} = \{\sigma | t\sigma =_{\mathcal{E}} t' \wedge Dom(()\sigma) \subseteq vars(t)\}$ for $t' \in L$. By Definition 3, $C_{t'}$ is not empty. So, let f be a choice function that, given a set $C_{t'}$, chooses a substitution σ from $C_{t'}$. Now, let $C = \bigcup_{t' \in L}\{f(C_{t'})\}$. Again, since L is finite, C is finite.

Let π be a substitution. Since L is complete, there is a term $t' \in L$, and a substitution τ such that $t\pi \downarrow_R = t'\tau$. Therefore, $t'\tau$ must be in normal form, so t' must also be in normal form. In addition, by the construction of $C_{t'}$, we have $t\sigma =_{\mathcal{E}} t'$, where $\sigma = f(C_{t'})$. Therefore, by the Church-Rosser property, and the fact that t' is in normal form, $t\sigma \downarrow_R = t'$. Therefore, we have $t\pi\downarrow_R = t'\tau = t\sigma \downarrow_R \tau$, and so C is complete. □

Ciobâcă uses the following example in [3] to show that the claim that a term t has a strongly complete finite set of variants is strictly stronger than the claim that t has a complete finite set of Ciob variants.

Example 4.1. Consider the equational theory $(\Sigma, \emptyset, R)$, where $\Sigma = ((\{S\}, \emptyset), \{a :\rightarrow S, f : S \rightarrow S, g : S \rightarrow S\}$, and $R = \{f(g(x)) = f(x)\}$.

In [3], Ciobâcă demonstrates that the set $\{id\}$, where id is the identity substitution, constitutes a complete set of Ciob-variants. Therefore, $(\Sigma, \emptyset, R)$ has the Ciob FVP (and by Theorem 2 the C-LD-FVP).

However, consider the term $f(x)$, and the substitution π_i where $\pi_i = \{x \mapsto g^i(y)\}$ for $i \in \mathbb{N}$. Any strongly complete set of variants in the sense of Ciobâcă must contain Ciob-variants of the form $\{x \mapsto g^i(z)\}$ for all $i \in \mathbb{N}$. Therefore, the set of strongly complete variants must be infinite, and so $(\Sigma, \emptyset, R)$ does not have the Ciob-SFVP even though it has the Ciob(C-LD)-FVP.

5. Comon-Lundh, Delaune and Escobar et al. Variants

In this section, we show that the EMS-FVP property implies the C-LD-FVP. The converse does not hold however, as will be shown in the following section, in which we will see that when $B = \emptyset$ the EMS-FVP is equivalent to the Ciob-SFVP.

Theorem 3. *Let (Σ, B, R) be an equational decomposition of $(\Sigma, \mathcal{E})$, where Σ contains a single sort. Let $\geq$ be $\rightarrow_{R,B}$ (and so $\min_{\geq} = \_\downarrow_{R,B}$). Then, for any Σ-term t, if t has a finite complete set of EMS variants, then t has a finite complete set of C-LD variants.*

Proof. Let S be a finite complete set of EMS-variants. Let $L = \{t' | \exists\theta, (t', \theta) \in V\}$. Note that since S is finite, L must be finite.

Let π be a substitution. Since S is complete, there exists $(t', \theta) \in S$ and a substitution τ s.t. $(t\pi) \downarrow_{R,B} =_B t'\tau$. Therefore, since $t' \in L$ by the definition of L, L is a complete set of C-LD-variants. □

6. Escobar et. al. Variants and Ciobâcă's Strong Finite Variant Property

In this section, we show that the EMS-FVP and Ciob-SFVP are equivalent in the case without axioms (since the Ciob-SFVP only handles the case without axioms).

Theorem 4. *Let $(\Sigma, \emptyset, R)$ be an equational decomposition of $(\Sigma, \mathcal{E})$, where Σ has a single sort and no subsorts. Let t be a Σ-term. Then, t has a finite complete set of EMS-variants, iff t has a finite strongly complete set of Ciob-variants.*

Proof. Let S be a complete set of EMS-variants of t. Let $C = \{\theta | (t', \theta) \in S\}$. Since S is finite, C must also be finite.

Let π be a substitution. Since S is complete, there exists $(t', \theta) \in S$ and a substitution τ such that t' is R-irreducible, $t\pi \downarrow_R = t'\tau$, and $\pi \downarrow_R |_{var(t)} = \theta\tau|_{var(t)}$. By construction, $\theta \in C$. Furthermore, observe that by definition, $t\theta \to^!_R t'$. Therefore, $t\pi \downarrow_R = (t\theta \downarrow_R)\tau$, and so C is a strongly complete set of Ciob-variants.

Now, let C be a strongly complete set of Ciob-variants. Let $S = \{(t\theta \downarrow_R, \theta \downarrow_R)|\theta \in C\}$.

Let π be a substitution. Since C is strongly complete, there exists $\theta \in C$, and a substitution τ such that $t\pi \downarrow_R = (t\theta \downarrow_R)\tau$, and $\pi \downarrow_R |_{var(t)} = (\theta \downarrow_R \tau)|_{var(t)}$. By construction, $(t\theta \downarrow_R, \theta \downarrow_R) \in S$. By the completeness of C we have τ such that $t\pi \downarrow_R = (t\theta \downarrow_R)\tau$, and $(\pi \downarrow_R |_{var(t)} = (\theta \downarrow_R)\tau|_{var(t)}$. So, S is a complete set of EMS-variants. □

7. Boundedness and Free Function Symbols

In this section, we investigate the relationship between boundedness and the various notions of variants. In particular, we investigate an apparent contradiction between Example 4.1 and a theorem given in [2] relating the Finite Variant Property and Boundedness.

First, a review of the definition of boundedness, from [2].

Definition 11 (Boundedness)**.** [2] Let $\mathcal{R} = (\Sigma, B, R)$ be an equational decomposition of equational theory $(\Sigma, \mathcal{E})$. Then, $\mathcal{R}$ satisfies the boundedness property iff for every Σ-term t, there is a number n such that for every R, B-normalized substitution σ, $t\sigma$ can be normalized in at most n-steps. Formally,

$$\forall t, \exists n, \forall \sigma . t(\sigma{\downarrow}) \overset{\leq n}{\to}_{R,B} (t\sigma){\downarrow}$$

Comon-Lundh and Delaune then proved the following key result in [2]:

Theorem 5. [2] *Let (Σ, B, R) be an equational decomposition of equational theory $(\Sigma, \mathcal{E})$. Then, (Σ, B, R) satisfies the boundedness property iff (Σ, B, R) has the finite variant property in the sense of C-LD.*

However, a careful study of Example 4.1 reveals that although it has the Ciob-FVP (and therefore the C-LD-FVP), it is not bounded:

Consider the term $f(x)$, and suppose $(\Sigma, \emptyset, E)$ has the boundedness property. Then, $f(x)$ has some bound, n. However, consider the term $f(g^{n+1}(y))$. The substitution $\{x \mapsto g^{n+1}(y)\}$ is normalized, yet it takes $n+1$ steps to normalize $f(g^{n+1}(y))$.

This merits some concern. A careful study of the proof in [2] reveals that it depends on the following lemma:

Lemma 1. *Let (Σ, B, E) be a decomposition of $(\Sigma, \mathcal{E})$. Then, (Σ, B, E) satisfies the finite variant property iff for every term t, there is a finite set of substitutions of t $\Theta(t)$ such that*

$$\forall \sigma, \exists \theta \in \Theta(t), \exists \tau . \sigma{\downarrow} =_B \theta\tau \wedge (t\sigma){\downarrow} =_B (t\theta){\downarrow}\tau$$

This lemma introduces the restrictions on the variant substitutions that are so pivotal in making both the EMS-FVP and Ciob-SFVP stronger than the C-LD-FVP. In order to prove this lemma, Comon-Lundh and Delaune add a free binary function symbol, $< \_, \_ >$ to Σ.

In other words, C-LD variants are only equivalent to boundedness if one is allowed to expand the signature. In addition to being an interesting subtlety, it also sheds some light on the modularity of the finite variant propery. Essentially what is being done in Lemma 1 is the union of (Σ, B, R) with the theory $(((S, \emptyset), \{< \_, \_ >\colon\ S\ S \to S\}), \emptyset, \emptyset)$, where S is the single sort in Σ. In the case of the theory in Example 4.1, as Ciobâcă demonstrates in [3], this is enough for the theory to lose the C-LD-FVP.

In addition, it should be observed that since Lemma 1 is built into the definitions of the strong finite variant property and the Escobar finite variant property, the argument used in [2] to prove Theorem 5 holds for the Ciob-SFVP and EMS-FVP even when Σ is fixed.

This gives us the following correct statement of Theorem 5:

Theorem 6. *Let (Σ, B, R) be a decomposition of order-sorted equational theory $(\Sigma, \mathcal{E})$. Then, (Σ, B, R) has the EMS-FVP iff (Σ, B, R) is bounded.*

This with Theorem 4 yield the following corollary:

Corollary 1. *Let $(\Sigma, \emptyset, R)$ be a decomposition of order-sorted equational theory $(\Sigma, \mathcal{E})$, where Σ contains a single sort and no subsorts. Then, $(\Sigma, \emptyset, R)$ has the Ciob-SFVP iff $(\Sigma, \emptyset, R)$ is bounded.*

In short, Ciob-SFVP and EMS-FVP are strictly stronger than the C-LD-FVP only if Σ is fixed. If we allow the introduction of arbitrary free function symbols to Σ, then the three notions become equivalent in the case without axioms, and the EMS-FVP and C-LD-FVP become equivalent in the case with axioms.

8. Checking the Escobar et al. Finite Variant Property

Now that the relationship between the EMS-FVP and boundedness has been fully defined in Theorem 6, we can explore a key result of this report: a simple semi-decision procedure for checking if an equational theory $(\Sigma, \mathcal{E})$ has the EMS-FVP.

Theorem 7. *Let (Σ, B, R) be a decomposition of an equational theory $(\Sigma, \mathcal{E})$. Then, $(\Sigma, \mathcal{E})$ has the EMS-FVP under the decomposition (Σ, B, R) if and only if for each*

$f : s_1, \ldots, s_n \to s \in \Sigma$, the term $f(x_1 : s_1, \ldots, x_n : s_n)$, where $x_1, \ldots, x_n$ are distinct variables, has a finite set of most general EMS-variants.

Proof. ($\Rightarrow$) is trivial. To see the ($\Leftarrow$) direction, we reason by structural induction, showing that each Σ-term satisfies the boundedness property. The base case is that of a constant a of sort s, which trivially satisfies the boundedness property by choosing the length of any rewrite sequence $a \to_{R,B}{}^{!} a \downarrow_{R,B}$.

For the induction step, assume that in the term $f(t_1, \ldots, t_n)$, the terms t_i have bounds n_i. Let $f : s_1 \ldots s_n \to s$ be the type of f in the above term, and let $\{(v_1, \theta_1), \ldots, (v_m, \theta_m)\}$ be the set of most general EMS-variants of $f(x_1 : s_1, \ldots, x_n : s_n)$. Let k_i be the length of a rewrite sequence $f(x_1, \ldots, x_n)\theta_i \to_{R,B}{}^{!} v_i$, and let $k = \max(\{k_i | 1 \leq i \leq m\})$. We claim that any instance $f(t_1, \ldots, t_n)\gamma$ for any γ in R, B-normal form can be reduced in less than $k + \Sigma_{i=1}^{n} n_i$ steps. Indeed, by the induction hypothesies we have a reduction $f(t_1, \ldots, t_n)\gamma \to_{R,B}{}^{*} f(t_1\gamma \downarrow_{R,B}, \ldots, t_n\gamma \downarrow_{R,B})$ in $\Sigma_{i=1}^{n} n_i$ steps giving us an R, B-normalized substitution $\alpha = \{x_1 : s_1 \mapsto t_1\gamma \downarrow_{R,B}, \ldots, x_n : s_n \mapsto t_n\gamma \downarrow_{R,B}\}$.

Therefore, $(f(t_1\gamma \downarrow_{R,B}, \ldots, t_n\gamma \downarrow_{R,B}) \downarrow_{R,B}, \alpha)$ is an EMS-variant of $f(x_1, \ldots, x_n)$, and thus there exists a variant (v_j, θ_j) and an R, B-normalized substitution β such that $\theta; \beta =_B \alpha$. This means that there is an instantiated rewrite sequence

$$f(x_1, \ldots, x_n)\theta \to_{R,B}{}^{!} v_i\beta$$

with n_i rewrite steps and $v_i\beta$ is up to B-equality a canonical form of $f(t_1, \ldots, t_n)\gamma$, since we have $v_i\beta =_B f(t_1\gamma \downarrow_{R,B}, \ldots, t_n\gamma \downarrow_{R,B}) \downarrow_{R,B}$, and of course $f(t_1, \ldots, t_n)\gamma \downarrow_{R,B} = f(t_1\gamma \downarrow_{R,B}, \ldots, t_n\gamma \downarrow_{R,B}) \downarrow_{R,B}$.

Therefore, the length of the sequence $f(t_1, \ldots, t_n)\gamma \to_{R,B}{}^{*} f(t_1\gamma \downarrow_{R,B}, \ldots, t_n\gamma \downarrow_{R,B}) \to_{R,B}{}^{*} f(t_1 \ldots, t_n)\gamma \downarrow_{R,B}$ is $\Sigma_{i=1}^{n} n_i + k_j \leq \Sigma_{i=1}^{n} n_i + k$ steps, as desired. □

Example 8.1. To illustrate this result, we will prove that the theory of Abelian Groups has the EMS-FVP. In [2], Comon-Lundh and Delaune proved that the theory of Abelian Groups has the C-LD-FVP under a well-known decomposition originally proposed by Lankford, written below in the declarative programming language Maude (for details about Maude, see [1]).

```
fmod ABELIAN-GROUP is
       sort Element  .
       op _+_ : Element Element -> Element
            [assoc comm prec 30] .
       op -_ : Element -> Element  [prec 20] .
       op 0 : -> Element .
       vars X Y Z : Element  .
       eq X + 0           = X             .
       eq X + - X         = 0             .
       eq X + - X + Y     = Y             .
       eq - - X           = X             .
       eq - 0             = 0             .
       eq - X + - Y       = -(X + Y)      .
       eq -(X + Y) + Y    = - X           .
       eq -(- X + Y)      = X + - Y       .
       eq - X + - Y + Z   = -(X + Y) + Z .
       eq -(X + Y) + Y + Z = - X + Z      .
```

```
      endfm
```

Well-formed Maude functional modules (deliminated by the `fmod` and `endfm` keywords), are convergent rewrite theories. Therefore, the above functional module is an equational decomposition of the theory of Abelian Groups.

Sorts are declared with the `sort` keyword.

Function symbols (operators) are declared using the `op` keyword. Operators may be defined using mixfix syntax, with underscores denoting the position of each of the operator's arguments. The type of a function is denoted with a typewriter approximation of the standard notation for function declaration, e.g. `_+_` takes two arguments of sort `Element`, and returns a term of sort `Element`. An operator with no arguments (`0`) is a constant.

It is also possible to annotate operators with certain built-in axioms (defining the axioms B). In this example, the `_+_` operator is denoted with the `assoc` and `comm` keywords, making `_+_` associative and commutative.

The `prec 20` and `prec 30` attributes provide parsing information. In particular, these two attributes say that `-_` binds more tightly than `_+_`.

The `vars` keyword is used to denote variable declarations. So, `X`, `Y`, and `Z` are variables of sort `Element`.

The equations denoted with the `eq` keyword are oriented from left to right, and define the set of rewrite rules R.

Furthermore, an experimental version of the Maude interactive environment has a command `get variants`, that generates all the variants of a term. Therefore, in order to prove that the theory of Abelian Groups has the finite variant property, we need only call the following two commands in the Maude interpreter(note that since 0 is a constant, and in normal form, it only has one variant: $\{(0, id)\}$):

(1) `Maude> get variants X + Y .`
(2) `Maude> get variants - X .`

If both commands terminate and generate a finite number of EMS-variants, then the theory of Abelian Groups will have the EMS-FVP.

Running these commands yields 47 EMS-variants for $x + y$ and 4 EMS-variants for $-x$.

Five of the generated EMS-variants of $x+y$ are: $\{(x+y, id), (y, \{x \mapsto 0\}), (0, \{y \mapsto -x\}), (x_1 + y_1, \{x \mapsto x_2 + x_1, y \mapsto y_1 + -x_2\}), (x_1, \{x \mapsto x_1 + x_2, y \mapsto -x_2\})\}$

The computed four most general EMS-variants of $-x$ are: $\{(-x, id), (x_1, \{x \mapsto -x_1\}), (0, \{x \mapsto 0\}), (x_1 + -x_2, \{x \mapsto x_2 + -x_1\})\}$

In other words, in order to prove that a given equational decomposition (Σ, B, R) of $(\Sigma, \mathcal{E})$ has the EMS-FVP, one merely needs to compute the most general EMS-variants of the terms $f(x_1 : s_1, \ldots, x_n : s_n)$ for each $f : s_1 \ldots s_n \to s \in \Sigma$, where $x_1, \ldots, x_n$ are variables. Of course, if (Σ, B, R) does not have the finite variant property, then for at least one symbol f, the variant computation will not terminate. One method of computing variants is by using folding variant narrowing, discussed in detail in [6].

9. Conclusion

Since their introduction by Comon-Lundh and Delaune in [2], the notion of variants has been investigated in several different contexts. This has led to several different formulations of variants that on the surface appear to be equivalent: the original Comon-Lundh, and Delaune formulation [2], a formulation by Ciobâcă [3], and a formulation by Escobar, Meseguer, and Sasse [4]. Unfortunately, there is a certain fragility in the Comon-Lundh, and Delaune definition of variants that Ciobâcă demonstrated. In particular, there are situations in which a theory has the finite variant property, yet does not possess the boundedness property.

In this paper, we have studied the relationships between these three notions in terms of their respective finite variant properties. In the case without axioms, the Ciob-FVP and C-LD-FVP are equivalent. In the case without axioms and a fixed signature, the Ciob-SFVP is strictly stronger than the C-LD-FVP. However, when one is allowed to expand the signature, the two become equivalent. In the case without axioms, the strong finite variant property and the EMS-FVP are equivalent. In the case with a fixed signature, the EMS-FVP is strictly stronger than the C-LD-FVP. When one is allowed to expand the signature, the EMS-FVP and C-LD-FVP are equivalent.

We also highlighted what causes the relationship between boundedness and the C-LD-FVP to break down. In particular, in order for the equivalence for the C-LD-FVP and boundedness to hold, it must be the case that an equational theory (Σ, B, R) has the finite variant property if and only if for every term t there is a finite set of substitutions, $\Theta(t)$ such that the normalized form of every substitution σ decomposes into some $\theta \in \Theta(t)$ and some other substitution τ. Furthermore, the normal form of $t\sigma$ must be equal modulo B to the normal form of $t\theta$ instantiated with τ. For C-LD-variants, this property only holds if one is allowed to add a free binary function symbol to Σ. However, because this property is an integral part of the definition of both the EMS-FVP, and the Ciob-SFVP, neither requires expanding Σ.

Therefore, the equivalence between boundedness and the EMS-FVP holds. We then used this equivalence to prove a key result of this report, which provides a simple method for checking for the finite variant property. Given an equational decomposition of (Σ, B, R) of equational theory $(\Sigma, \mathcal{E})$, one can prove that (Σ, B, R) has the finite variant property by showing that the term $f(x_1 : s_1, \ldots, x_n : s_n)$ has a finite set of most general variants for every operator $f : s_1 \ldots s_n \to s \in \Sigma$.

References

[1] M. Clavel, F. Durán, S. Eker, P. Lincoln, N. Martıi-Oliet, J. Meseguer, and C. L. Talcott. All about maude - a high-performance logical framework. In *Lecture Notes in Computer Science*, volume 4350. Springer, 2007.

[2] H. Comon-Lundh and S. Delaune. The finite variant property: How to get rid of some algebraic properties. In J. Giesl, editor, *Lecture Notes in Computer Science*, volume 3467, pages 294–307. Springer, 2005.

[3] Ştefan Ciobâcă. *Verification of Composition of Security Protocols with Applications to Electronic Voting*. PhD thesis, ENS Cachan, December 2011.

[4] S. Escobar, J. Meseguer, and R. Sasse. Effectively checking the finite variant property. In *Lecture Notes in Computer Science*, volume 5117, pages 79–93. Springer, 2008.

[5] S. Escobar, J. Meseguer, and R. Sasse. Variant narrowing and equational unification. In *Electronic Notes Theoretical Computer Science*, volume 238, pages 103–119, 2009.

[6] S. Escobar, R. Sasse, and J. Meseguer. Folding variant narrowing and optimal variant termination. *Journal of Logic and Algebraic Programming*, 81:898–928, 2009.

[7] J. Giesl and D. Kapur. Dependency pairs for equational rewriting. In A. Middeldorp, editor, *Lecture Notes in Computer Science*, volume 2051, pages 93–108. Springer, 2001.

[8] J.-P. Jouannaud, C. Kirchner, and H. Kirchner. Incremental construction of unification algorithms in equational theories. In J. Díaz, editor, *Lecture Notes in Computer Science*, volume 154, pages 361–373. Springer, 1983.

[9] J.-P. Jouannaud and H. Kirchner. Completion of a set of rules modulo a set of equations. *SIAM Journal of Computing*, 15:1155–1194, 1986.

[10] J. Meseguer. Conditional rewriting as a unified model of concurrency. *Theoretical Computer Science*, 96:73–155, 1992.

[11] J. Meseguer. Membership algebra as a logical framework for equational specification. In F. Parisi-Presicce, editor, *Lecture Notes in Computer Science*, volume 1376, pages 18–61. Springer, 1997.

[12] G. E. Peterson and M. E. Stickel. Complete sets of reductions for some equational theories. *Journal of the ACM*, 28:233–264, 1981.

[13] TeReSe, editor. *Term Rewriting Systems*. Cambridge University Press, Cambridge, 2003.

[14] P. Viry. Equational rules for rewriting logic. *Theoretical Computer Science*, 285:487–517, 2002.