32nd International Workshop on Unification (UNIF 2018), Informal Proceedings
Oxford, 7th July, 2018

Knowledge Problems in Equational Extensions of Subterm Convergent Theories

(Extended Abstract)

Serdar Erbatur[1], Andrew M. Marshall[2], and Christophe Ringeissen[3]*

[1] IMDEA Software Institute (Spain)
serdar.erbatur@imdea.org
[2] University of Mary Washington (USA)
marshall@umw.edu
[3] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
Christophe.Ringeissen@loria.fr

Abstract

We study decision procedures for two knowledge problems critical to the verification of security protocols, namely the intruder deduction and the static equivalence problems. These problems can be related to particular forms of context matching and context unification. Both problems are defined with respect to an equational theory and are known to be decidable when the equational theory is given by a subterm convergent term rewrite system. In this note we extend this to consider a subterm convergent equational term rewrite system defined modulo an equational theory, like Commutativity or Associativity-Commutativity. We show that for certain classes of such equational theories, namely the shallow classes, the two knowledge problems remain decidable.

1 Introduction

Verifying the security of protocols requires the development of specific decision procedures to reason about the knowledge of an intruder. Two important measures of this knowledge are *(intruder) deduction* [17,18] and *static equivalence* [2]. The deduction problem is the question of whether an intruder, given his deductive capability and a sequence of messages representing their knowledge, can obtain some secret. This is a critical measure of the capability of the protocol to maintain secrets. Deducibility is needed for many questions about the security of protocols. However, there are some questions for which we need to be able to decide more than deducibility. For some protocols, in addition to deducibility, we would like to determine whether an intruder can distinguish between different runs of the protocol. For example, in protocols which attempt to transmit encrypted votes we would like to know if, to the attacker, two different votes are indistinguishable. Static equivalence measures this property.

*This work has received funding from the European Research Council (ERC) under the H2020 research and innovation program (grant agreement No 645865-SPOOC).

Much work has gone into investigating and developing decision procedures for the deduction and the static equivalence problems [2,6,8,10,12]. In this work the security protocols are often represented by equational theories with the theories of interest usually defined as unions of several simpler sub-theories. In this paper, we focus on decision procedures for the deduction problem and the static equivalence problem in equational theories $E = E_1 \cup E_2$ where E_1 and E_2 are possibly non-disjoint. Until now, the following scenarios have been successfully investigated: E_1 is given by a subterm convergent term rewrite system, and E_2 is empty [2]; E_1 and E_2 are disjoint [9]; E_1 and E_2 share only constructors [12]. In this paper, we consider the case where E_1 is given by a term rewrite system which is both subterm and convergent modulo E_2. We then show that the methods of [2] can be extended to subterm convergent rewrite systems for a significant class of E_2 theories.

2 Preliminaries

We assume the reader is familiar with equational unification and term rewrite systems. We review some critical definitions below but a more complete overview can be found in [5].

A finite convergent *term rewrite system* (TRS) R is said to be *subterm convergent* if for any $l \to r \in R$, r is either a strict subterm of l or a ground term. An equational theory is *subterm convergent* if it is presented by a subterm convergent TRS.

The size of a term t is denoted by $|t|$ and defined in the usual way as follows: $|f(t_1, \ldots, t_n)| = 1+\Sigma_{i=1}^{n}|t_i|$ if f is a n-ary function symbol with $n \geq 1$, $|c| = 1$ if c is a constant, and $|x| = 0$ if x is a variable. The size of a TRS R is denoted by $|R|$ and defined as follows: $|R| = \max_{\{l \to r \in R\}} |l|$. Since a variable cannot occur as the left-hand side of any rule in R, we have that $|R| \geq 1$ for any non-empty TRS R. When R is empty, we define $|R| = 1$.

Equational Theories. Let us introduce the different classes of theories considered in the paper. An equational theory E is *finite* if for each term t, there are only finitely many terms s such that $t =_E s$. Matching in finite theories is finitary. A sufficient condition to get a finite theory is to assume that E is permutative. An equational theory E is *permutative* if for each axiom $l = r$ in E, l and r contain the same symbols with the same number of occurrences. Well-known theories such as Associativity (A), Commutativity (C), and Associativity-Commutativity (AC) are permutative theories. Unification in permutative theories is undecidable in general [19].

A theory E is *syntactic* if it has a finite *resolvent presentation* S, that is a finite set of equational axioms S such that each equality $t =_E u$ has an equational proof $t \leftrightarrow_S^* u$ with at most one step $\leftrightarrow_S$ applied at the root position. A theory E is *shallow* if variables can only occur at a depth at most 1 in axioms of E. Shallow theories are syntactic theories for which unification is finitary [7]. The theory AC is permutative and syntactic, while C is permutative and shallow.

Notions of Knowledge. The applied pi calculus and frames are used to model attacker knowledge [3]. In this model, the set of messages or terms which the attacker knows, and which could have been obtained from observing one or more protocol sessions, are the set of terms in $Ran(\sigma)$ of the frame $\phi = \nu\tilde{n}.\sigma$, where σ is a substitution ranging over ground terms. We also need to model cryptographic concepts such as nonces, keys, and publicly known values. We do this by using names, which are essentially free constants. Here also, we need to track the names which the attacker knows, such as public values, and the names which the attacker does not know a priori, such as freshly generated nonces. $\tilde{n}$ consists of a finite set of restricted names,

these names represent freshly generated names which remain secret from the attacker. The set of names occurring in a term t is denoted by $fn(t)$.

Definition 1 (Deduction Problem [2]). *Let $\phi = \nu\tilde{n}.\sigma$ be a frame, and t a ground term. We say that t is* deduced from ϕ modulo E, *denoted by $\phi \vdash_E t$, if there exists a term s such that $s\sigma =_E t$ and $fn(s) \cap \tilde{n} = \emptyset$. The term s is called a* recipe *of t in ϕ modulo E.*

Another form of knowledge is the ability to tell if two frames are *statically equivalent* modulo E, sometimes also called *indistiguishability*.

Definition 2 (Static Equivalence [2]). *Two terms s and t are* equal *in a frame $\phi = \nu\tilde{n}.\sigma$ modulo an equational theory E, denoted $(s =_E t)\phi$, iff $s\sigma =_E t\sigma$, and $\tilde{n} \cap (fn(s) \cup fn(t)) = \emptyset$. Two frames $\phi = \nu\tilde{n}.\sigma$ and $\psi = \nu\tilde{n}.\tau$ are* statically equivalent modulo E, *denoted as $\phi \approx_E \psi$, if $Dom(\sigma) = Dom(\tau)$ and for all terms s and t, we have $(s =_E t)\phi$ iff $(s =_E t)\psi$.*

Both deduction and static equivalence are known to be decidable in subterm convergent theories [2]. In the following, we lift this result to term rewrite systems that are subterm convergent modulo some equational theory.

3 Subterm Equational Convergent TRS

Consider $(\Sigma, E) = (\Sigma_1 \cup \Sigma_2, R_1 \cup E_2)$ where $(\Sigma_1 \cup \Sigma_2, R_1)$ is a TRS modulo a finite theory (Σ_2, E_2) (for instance $\Sigma_2 = \{+\}$ and $E_2 = AC(+)$). The rewrite relation of R_1 modulo E_2 is defined as usual: $s \to_{R_1,E_2} t$ if there exist some position p in s, some rule $l \to r \in R_1$ and a substitution μ such that $s|_p =_{E_2} l\mu$ and $t = s[r\mu]_p$. We assume that $\to_{R_1,E_2}$ is convergent modulo E_2 [15]. This implies the uniqueness of normal forms modulo E_2 and the decidability of the word problem modulo E: for any terms s and t, we have $s =_E t$ iff $(s \downarrow_{R_1,E_2}) =_{E_2} (t \downarrow_{R_1,E_2})$. In the following, we say that a term or a substitution is normalized if it is normalized w.r.t $\to_{R_1,E_2}$. A frame $\phi = \nu\tilde{n}.\sigma$ is said to be normalized if σ is normalized.

Definition 3. *Let Σ_1 and Σ_2 be two disjoint signatures, and (Σ_2, E_2) a finite theory. A* subterm E_2-convergent *TRS $(\Sigma_1 \cup \Sigma_2, R_1)$ is a TRS such that $\to_{R_1,E_2}$ is convergent modulo E_2 and for any $l \to r$ in R_1, l is not Σ_2-rooted and r is a strict subterm of l or a ground term.*

Example 1. *The following TRSs are subterm $AC(+)$-convergent:*

$\{occ(x + k, k) \to ok\}$	$\{rm(x + k, k) \to x\}$
$\{dec(enc(x, k + y), k) \to x\}$	$\{dec(enc(x, k), k + y) \to x\}$

In the case of subterm convergent TRSs (modulo the empty theory), the decision procedure for the deduction problem computes deducible terms among the set of subterms occurring in the frame. When considering a non-empty theory E_2, we have to introduce an extended notion of subterm to capture the fact that matching modulo E_2 is now performed when applying a rewrite step modulo E_2.

In the rest of this section we assume that E_2 is both permutative and syntactic. While this may seem somewhat restrictive it allows for the consideration of theories such as AC and C which are found in a large number of security protocols. Both AC and C are indeed syntactic theories [16].

Given a term t, $St(t)$ is the finite set of terms in t inductively defined as follows:

$$St(t) = \{t' \mid t' =_{E_2} t\} \cup \left\{ t' \;\middle|\; \begin{array}{l} t' \in St(x_i\sigma), f(x_1, \ldots, x_m)\sigma =_{E_2} t, f \in \Sigma_1 \cup \Sigma_2 \\ x_1, \ldots, x_m \text{ are pairwise disjoint variables} \end{array} \right\}$$

This definition is well-founded since E_2 is permutative. There exists a mutation-based E_2-matching algorithm [11] since E_2 is syntactic, and so $St(t)$ is computable.

For a set of terms T, $St(T) = \bigcup_{t \in T} St(t)$, and for a substitution σ, $St(\sigma) = St(Ran(\sigma))$.

Proposition 1. *For any terms t, t', $t =_{E_2} t'$ implies $St(t) = St(t')$, and for any position p in t, $St(t|_p) \subseteq St(t)$.*

The following result states that we cannot generate a new term outside $St(t)$ by rewriting terms in $St(t)$ (except the ground right-hand sides of R_1).

Lemma 1. *If $l\sigma =_{E_2} t$, then for any position p of l, $(l|_p)\sigma \in St(t)$.*

Proof. By structural induction on l.

If l is a variable, this is trivial since the only possible position is ϵ and $l|_\epsilon = l$.

Assume l is a term $f(l_1, \ldots, l_m)$ and σ is a substitution such that $f(l_1, \ldots, l_m)\sigma =_{E_2} t$.

If there is an equational step at the root position, then there exist some terms $g_1, \ldots, g_m$ such that $l_1\sigma =_{E_2} g_1, \ldots, l_m\sigma =_{E_2} g_m$ and $f(g_1, \ldots, g_m) =_{E_2} t$. By definition of $St(t)$ and Proposition 1, the terms $g_1, \ldots, g_m$ are in $St(t)$, and so $l_1\sigma, \ldots, l_m\sigma \in St(t)$.

If there is no equational step at the root position, then t is of the form $f(t_1, \ldots, t_m)$ and $l_1\sigma =_{E_2} t_1, \ldots, l_m\sigma =_{E_2} t_m$. By definition of $St(t)$ and Proposition 1, the terms $t_1, \ldots, t_m$ are in $St(t)$, and so $l_1\sigma, \ldots, l_m\sigma \in St(t)$. □

4 Decision Procedures

From now on, we assume that E_2 is a shallow permutative theory, e.g., Commutativity.

Deduction. The decision procedure for the deduction problem requires the computation of some finite deducible terms defining the so-called *completion* of a given frame.

Definition 4. *Let $\phi = \nu\tilde{n}.\sigma$ be a normalized frame. The set of* local deducible terms in ϕ *is the smallest set D such that:*

- $Ran(\sigma) \subseteq D$,
- *if $t_1, \ldots, t_n \in D$ and $f(t_1, \ldots, t_n) \in St(\sigma)$ then $f(t_1, \ldots, t_n) \in D$,*
- *if $t \in D$, $t' \in St(\sigma)$, $t =_{E_2} t'$, then $t' \in D$,*
- *if there is a root reduction $s[\bar{r}] \rightarrow^{\epsilon}_{R_1, E_2} t$ where $|s| \leq |R_1|$, $fn(s) \cap \tilde{n} = \emptyset$, $\bar{r} \in D$ and $t \in St(\sigma)$, then $t \in D$.*

Let $\sigma_* = \sigma\{\chi_u \mapsto u \mid u \in D \backslash Ran(\sigma)\}$ where χ_u is a fresh variable. The frame $\phi_* = \nu\tilde{n}.\sigma_*$ is called the completion of ϕ with respect to R_1. *The* recipe substitution *of ϕ is $\zeta_\phi = \{\chi_u \mapsto \zeta_u \mid u \in D \backslash Ran(\sigma)\}$ where ζ_u denotes an arbitrary recipe of u w.r.t. ϕ.*

The decision procedure is based on the following reduction lemma, using the facts that the completion is computable and the deduction problem is decidable in the empty equational theory.

Lemma 2. *Let $E = R_1 \cup E_2$ where R_1 is any subterm E_2-convergent TRS and E_2 is any shallow permutative theory. For any normalized frame ϕ and any normalized term t, we have that $\phi \vdash_E t$ if and only if $\phi_* \vdash t$.*

Static Equivalence. The decision procedure for the static equivalence is based on the computation of small equalities bounded by the size of R_1.

Definition 5. *Let $\phi = \nu\tilde{n}.\sigma$ be a normalized frame. The set $Eq(\phi)$ is the set of equalities $t\zeta_\phi = t'\zeta_\phi$ such that $(t\zeta_\phi)\sigma =_E (t'\zeta_\phi)\sigma$ where t, t' are Σ-terms, $(fn(t) \cup fn(t')) \cap \tilde{n} = \emptyset$, $|t|, |t'| \leq |R_1|$. Given any frame $\psi = \nu\tilde{n}.\tau$, the fact that $t\tau =_E t'\tau$ for any $t = t' \in Eq(\phi)$ is denoted by $\psi \models Eq(\phi)$.*

To get a decision procedure, it remains to show that checking small equalities defined by Eq are sufficient to prove the static equivalence of the two input frames. Note that the check of each of these equalities is effective since the E-equality is decidable.

Lemma 3. *Let $E = R_1 \cup E_2$ where R_1 is any subterm E_2-convergent TRS and E_2 is any shallow permutative theory. For any normalized frames ϕ and ψ, we have that $\phi \approx_E \psi$ iff $\psi \models Eq(\phi)$ and $\phi \models Eq(\psi)$.*

Main result. According to the above reduction lemmas, we get the following result.

Theorem 1. *Let $E = R_1 \cup E_2$ where R_1 is any subterm E_2-convergent TRS and E_2 is any shallow permutative theory. Then, deduction and static equivalence are decidable in E.*

To prove both reduction lemmas (Lemma 2 and Lemma 3) and so Theorem 1, we reuse the same approach as in [1,2] by applying two technical lemmas.

The first lemma in the appendix of [1] can be generalized as follows.

Lemma 4. *Let $E = R_1 \cup E_2$ where R_1 is any subterm E_2-convergent TRS and E_2 is any shallow permutative theory. For any terms s and t satisfying the name restriction, if $s\phi_* =_{E_2} t\phi_*$ and $\psi \models Eq(\phi)$ then $(s\zeta_\phi)\psi =_E (t\zeta_\phi)\psi$.*

Then, the second lemma in the appendix of [1] is generalized in the following way.

Lemma 5. *Let $E = R_1 \cup E_2$ where R_1 is any subterm E_2-convergent TRS and E_2 is any shallow permutative theory. For any term s satisfying the name restriction and for any term t such that $s\phi_* \rightarrow_{R_1,E_2} t$, there exists a term u satisfying the name restriction such that $t =_{E_2} u\phi_*$ and for any frame ψ such that $\psi \models Eq(\phi)$, $(s\zeta_\phi)\psi =_E (u\zeta_\phi)\psi$.*

The proofs of Lemma 4 and Lemma 5 can be found in [13]. The assumption that E_2 is shallow permutative allows us to get simple proofs.

We are working on generalizing Theorem 1 to syntactic permutative theories E_2 like for instance Associativity-Commutativity. In this general case, the related reduction lemmas for the deduction problem and the static equivalence should be more complicated to express. Indeed, we may have to integrate a deduction procedure modulo E_2 in the construction of the completion, and a static equivalence procedure modulo E_2 to get a reduction lemma for static equivalence in $R_1 \cup E_2$.

5 Related Work and Conclusion

The intruder deduction problem corresponds to the general cap problem studied in [4]. Among other results, it is shown in [4] that the general cap problem is in NP for *dwindling* convergent rewrite systems, which are indeed subterm convergent theories. The NP procedure is given

by a saturation procedure used to complete the knowledge given by the input frame. In the conclusion of [4], the extension to AC-rewrite systems is mentioned as an interesting future work.

Currently we assume in Definition 3 that the Σ_2-symbols are constructors, i.e., not appearing at the root of the left-hand sides of the rewrite system. However, this appears to be more restrictive than needed. Indeed, it should be possible to remove this restriction and consider a more relaxed definition where the Σ_2-symbols are not necessarily constructors. This would allow us to solve the deduction and static equivalence problem in a larger class of theories. For example, we could then consider the theory of Abelian Pre-Group (APG) defined by the following C-convergent TRS:

$$R_{APG} = \{x * e \to x,\ x * i(x) \to e,\ i(i(x)) \to x,\ i(e) \to e\}$$

where $C = \{x * y = y * x\}$. In [20], $APG = R_{APG} \cup C$ was considered as an approximation to deal with unification in homomorphic encryption over Abelian groups. Theorem 1 would then allow us to also solve the problems of deduction and static-equivalence in APG.

The next step of our work is to go beyond the class of shallow permutative theories, in order to take into account a larger class including AC. Due to the potential interest of AC in protocol analysis, it is useful to be able to handle some AC-rewrite systems and to study the AC-extension of saturation procedures that have been developed for the intruder deduction problem, the static equivalence [6], and the static inclusion [14].

Another challenging problem is to investigate the equational extension of the combination procedure developed in [12] for the deduction and the static equivalence in unions of theories sharing absolutely free constructors. This would permit us to consider shared AC-constructors.

Acknowledgments: We would like to thank Véronique Cortier for helpful comments and discussions.

References

[1] Martín Abadi and Véronique Cortier. Deciding knowledge in security protocols under equational theories. Research Report RR-5169, INRIA, 2004.

[2] Martín Abadi and Véronique Cortier. Deciding knowledge in security protocols under equational theories. *Theor. Comput. Sci.*, 367(1-2):2–32, 2006.

[3] Martín Abadi and Cédric Fournet. Mobile values, new names, and secure communication. In *Proceedings of the 28th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL'01, pages 104–115, New York, NY, USA, 2001. ACM.

[4] Siva Anantharaman, Paliath Narendran, and Michaël Rusinowitch. Intruders with caps. In Franz Baader, editor, *Term Rewriting and Applications, 18th International Conference, RTA 2007, Paris, France, June 26-28, 2007, Proceedings*, volume 4533 of *Lecture Notes in Computer Science*, pages 20–35. Springer, 2007.

[5] Franz Baader and Tobias Nipkow. *Term rewriting and all that.* Cambridge University Press, New York, NY, USA, 1998.

[6] Mathieu Baudet, Véronique Cortier, and Stéphanie Delaune. YAPA: A generic tool for computing intruder knowledge. *ACM Trans. Comput. Log.*, 14(1):4, 2013.

[7] Hubert Comon, Marianne Haberstrau, and Jean-Pierre Jouannaud. Syntacticness, cycle-syntacticness, and shallow theories. *Inf. Comput.*, 111(1):154–191, 1994.

[8] Bruno Conchinha, David A. Basin, and Carlos Caleiro. FAST: an efficient decision procedure for deduction and static equivalence. In Manfred Schmidt-Schauß, editor, *Proceedings of RTA 2011,*

Novi Sad, Serbia, volume 10 of *LIPIcs*, pages 11–20. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011.

[9] Véronique Cortier and Stéphanie Delaune. Decidability and combination results for two notions of knowledge in security protocols. *Journal of Automated Reasoning*, 48(4):441–487, 2010.

[10] Ştefan Ciobâcă, Stéphanie Delaune, and Steve Kremer. Computing knowledge in security protocols under convergent equational theories. *J. Autom. Reasoning*, 48(2):219–262, 2012.

[11] Serdar Erbatur, Deepak Kapur, Andrew M. Marshall, Paliath Narendran, and Christophe Ringeissen. Unification and matching in hierarchical combinations of syntactic theories. In Carsten Lutz and Silvio Ranise, editors, *Frontiers of Combining Systems - 10th International Symposium, FroCoS 2015, Wroclaw, Poland. Proceedings*, volume 9322 of *LNCS*, pages 291–306. Springer, 2015.

[12] Serdar Erbatur, Andrew M. Marshall, and Christophe Ringeissen. Notions of knowledge in combinations of theories sharing constructors. In Leonardo de Moura, editor, *Automated Deduction - CADE 26 - 26th International Conference on Automated Deduction, Gothenburg, Sweden, Proceedings*, volume 10395 of *LNCS*, pages 60–76. Springer, 2017.

[13] Serdar Erbatur, Andrew M. Marshall, and Christophe Ringeissen. Computing knowledge in equational extensions of subterm convergent theories. Available at https://hal.inria.fr, 2018.

[14] Kimberly A. Gero. *Deciding Static Inclusion for Delta-strong and Omega Delta-strong Intruder Theories: Applications to Cryptographic Protocol Analysis.* PhD thesis, State University of New York aty Albany, 2015.

[15] Jean-Pierre Jouannaud and Hélène Kirchner. Completion of a set of rules modulo a set of equations. *SIAM J. Comput.*, 15(4):1155–1194, 1986.

[16] C. Kirchner and F. Klay. Syntactic theories and unification. In *Logic in Computer Science, 1990. LICS '90, Proceedings., Fifth Annual IEEE Symposium on Logic in Computer Science*, pages 270–277, Jun 1990.

[17] Jonathan Millen and Vitaly Shmatikov. Constraint solving for bounded-process cryptographic protocol analysis. In *Proceedings of the 8th ACM Conference on Computer and Communications Security*, CCS'01, pages 166–175, New York, NY, USA, 2001. ACM.

[18] Lawrence C. Paulson. The inductive approach to verifying cryptographic protocols. *Computer Security*, 6:85128, 1998.

[19] Manfred Schmidt-Schauß. Unification in permutative equational theories is undecidable. *J. Symb. Comput.*, 8(4):415–421, 1989.

[20] Fan Yang, Santiago Escobar, Catherine Meadows, José Meseguer, and Paliath Narendran. Theories of homomorphic encryption, unification, and the finite variant property. In *Proceedings of the 16th International Symposium on Principles and Practice of Declarative Programming*, PPDP '14, pages 123–133, New York, NY, USA, 2014. ACM.