Symbolic Protocol Analysis in Presence of a Homomorphism Operator and *Exclusive Or* ·

Stéphanie Delaune[1,2], Pascal Lafourcade[2,3], Denis Lugiez[3] and Ralf Treinen[2]

[1] France Télécom, Division R&D
[2] LSV, CNRS UMR 8643, ENS de Cachan & INRIA Futurs project SECSI
[3] LIF, Université Aix-Marseille1 & CNRS UMR 6166

Abstract. Security of a cryptographic protocol for a bounded number of sessions is usually expressed as a symbolic trace reachability problem. We show that symbolic trace reachability for *well-defined* protocols is decidable in presence of the *exclusive or* theory in combination with the homomorphism axiom. These theories allow us to model basic properties of important cryptographic operators.
This trace reachability problem can be expressed as a system of symbolic deducibility constraints for a certain inference system describing the capabilities of the attacker. One main step of our proof consists in reducing deducibility constraints to constraints for deducibility in one step of the inference system. This constraint system, in turn, can be expressed as a system of quadratic equations of a particular form over $\mathbb{Z}/2\mathbb{Z}[h]$, the ring of polynomials in one indeterminate over the finite field $\mathbb{Z}/2\mathbb{Z}$. We show that satisfiability of such systems is decidable.

1 Introduction

Cryptographic protocols are small programs designed to ensure secure communication via a network that may be controlled by an attacker. They involve a high level of concurrency and are difficult to analyze by hand. These programs are linear sequences of *receive* and *send* instructions on a public network. A *passive* attacker may only listen to messages, while an *active* attacker may also pretend to be a protocol participant and forge messages according to a certain set of *intruder capabilities*.

The problem of deciding whether a protocol preserves the confidentiality of a message under any active attack is known to be undecidable in general (*e.g.* [11]). Several decidability results have been obtained under the assumption that the number of role instances is bounded, among others NP-completeness due to Rusinowitch and Turuani [17]. The idea of their algorithm is to guess a symbolic trace in which the messages are represented by terms containing variables. This symbolic trace corresponds to a concrete execution trace if the variables can be instantiated in such a way that at every moment a message received by an agent can in fact be deduced by the intruder from the messages seen before. Hence, verifying security of a protocol amounts to a non-deterministic guessing of the symbolic trace plus the resolution of a system of

· This work has been partly supported by the RNTL project PROUVÉ 03V360 and the ACI-SI Rossignol.

deducibility constraints. This result [17], as many others (e.g., [15]), relies on the so-called *perfect cryptography assumption* which states that the cryptographic primitives (like encryption) are perfect and can be treated as black boxes. This assumption is unrealistic since some attacks exploit in a clever way the interaction between protocol rules and properties of cryptographic primitives. A more realistic approach is to take into account properties of the cryptographic primitives (see [4] for a survey). For the constraint based approach, this has been done for different equational theories [16, 8].

In this paper we study the equational theory ACUNh which is the combination of (h) the homomorphism axiom $h(x+y) = h(x) + h(y)$ with the *exclusive or* (ACUN) theory. These two equational theories model basic properties of important cryptographic primitives. Some protocols relying on these algebraic properties are described in [4]. *Exclusive or* is a basic building block in many symmetric encryption methods like DES or AES, or even used directly as an encryption method (Vernam encryption). Homomorphisms are ubiquitous in cryptography. For instance, the Wired Equivalent Privacy (WEP) protocol uses a checksum function C which has the homomorphism property over $+$, *i.e.* $C(x + y) = C(x) + C(y)$. Moreover, the homomorphism property over some binary operator appears in several encryption schemes (RSA, ElGamal ...) and is crucial in the field of electronic voting protocols [5]. Note that the recent result by Chevalier and Rusinowitch [2] for the combination of intruder theories can not be employed here to simply extend the known decidability result [1, 3] for ACUN since the theories ACUN and h share the symbol $+$. Furthermore, their result relies on a model which is different from ours in that it applies only to a restricted class of protocols.

Some results have already been obtained for the ACUNh theory [13, 6], but only for the case of a passive attacker. This algorithm for passive attacks is an important ingredient to the algorithm for active attacks developed in the present paper. Another important ingredient is ACUNh unification which has been shown decidable in [12]. However, for our procedure, we need to establish that unification in ACUNh is finitary, i.e. that every problem has a finite set of most general solutions. Our work is inspired by Millen and Shmatikov's approach [16] for the equational theory of Abelian groups. However, there are fundamental differences in the technical development.

Outline of the paper. We present our attacker model in Section 2, and the classes of constraint systems that we employ in our algorithm in Section 3. The proof of our main result (Theorem 1) proceeds in two steps: First we reduce satisfiability of deducibility constraints to satisfiability of constraints for one-step deducibility by a particular inference rule (Section 4). Second, we reduce satisfiability of these constraints to the satisfiability of a particular form of quadratic equations over the ring $\mathbb{Z}/2\mathbb{Z}[h]$, which we finally show to be decidable in Section 5 (satisfiability of quadratic equations over $\mathbb{Z}/2\mathbb{Z}[h]$, or for that matter $\mathbb{Z}$, is undecidable in general). Due to lack of space, proofs are omitted and can be found in [9].

2 Attacker Model

2.1 Inference System

The deduction capabilities of the intruder are formalized by the *Dolev-Yao model* [10]. We extend the intruder capabilities by equational reasoning modulo a given set E of equational axioms; we denote this intruder model by $\mathcal{I}_{\mathsf{DY+E}}$. In this paper, we consider the equational theory $\mathsf{E} = \mathsf{ACUNh}$ which consists of the well-known axioms of *exclusive or* in combination with a homomorphism symbol. More formally, ACUNh contains the following equations:

- Associativity, Commutativity (AC): $x + (y + z) = (x + y) + z$, $x + y = y + x$,
- Unit (U): $x + 0 = x$,
- Nilpotence (N): $x + x = 0$,
- homomorphism (h): $h(x + y) = h(x) + h(y)$.

We obtain the inference system described in Figure 1 where equational reasoning is taken into account through the normalization function $\downarrow$ associated to E. In the case of the ACUNh equational theory, the AC-convergent rewrite system is obtained by orienting from left to right the equations (U), (N), (h) and by adding the consequence $h(0) \rightarrow 0$ (see [13] for details). We omit the equality rule for AC and just work with equivalence classes modulo AC.

$$\text{Unpairing (UL)}\ \frac{T \vdash \langle u, v\rangle}{T \vdash u} \qquad \text{Compose (C)}\ \frac{T \vdash u_1 \ \ldots \ T \vdash u_n}{T \vdash f(u_1, \ldots, u_n)} \text{ with } f \in \mathcal{F} \smallsetminus \{+, h, 0\}$$

$$\text{Unpairing (UR)}\ \frac{T \vdash \langle u, v\rangle}{T \vdash v} \qquad \text{Context}(\mathsf{M_E})\ \frac{T \vdash u_1 \ \ldots \ T \vdash u_n}{T \vdash C[u_1, \ldots, u_n]\downarrow} \text{ with } C \text{ an } \mathsf{E}\text{-context}$$

$$\text{Decryption (D)}\ \frac{T \vdash \{u\}_v \quad T \vdash v}{T \vdash u}$$

Fig. 1. Dolev-Yao Model Extended with an Equational Theory: $\mathcal{I}_{\mathsf{DY+E}}$

The intended meaning of a *sequent* $T \vdash u$ is that the intruder is able to deduce the term $u \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ from the finite set of terms $T \subseteq \mathcal{T}(\mathcal{F}, \mathcal{X})$. As in the standard Dolev-Yao model, the intruder can compose new terms from known terms (C), he can decompose pairs (UL, UR), and he can decrypt ciphertexts provided that he can deduce the decryption key (D). Finally, the intruder may apply ($\mathsf{M_E}$) any E-context, *i.e.* term of the form $C[x_1, \ldots, x_n]$ with $C \in \mathcal{T}(\{0, +, h\}, \{x_1, \ldots, x_n\})$, to terms he already knows. Examples of instances of this rule are

$$\frac{T \vdash a + h(a) \qquad\qquad T \vdash b}{T \vdash a + h(h(h(a))) + h(b)}\ (\mathsf{M_E}) \qquad\qquad \frac{}{T \vdash 0}\ (\mathsf{M_E})$$

obtained with $C[x_1, x_2] = x_1 + h(x_1) + h(h(x_1)) + h(x_2)$, resp. $C[] = 0$.

The notation $h^n(t)$ represents the term t if $n = 0$ and $h(h^{n-1}(t))$ otherwise. Along this paper, we consider implicitly that terms are kept in normal form, *i.e.* we write u (resp. $u\sigma$) instead of $u\downarrow$ (resp. $u\sigma\downarrow$).

This deductive system is equivalent in deductive power to a variant of the system in which terms are not automatically normalized, but in which arbitrary equational proofs are allowed at any moment of the deduction (see [6, 13]). The inference system described in Figure 1 deals with symmetric encryption. However, it is not difficult to design a similar deduction system for asymmetric encryption and to extend the results of this paper to this new inference system.

2.2 Factors, Subterms

A term t is *standard* if and only if it is not of the form $f(t_1, \ldots, t_n)$ for some term $t_1, \ldots, t_n$ and some $f \in \{0, h, +\}$. In particular, every variable is a standard term.

Definition 1. *Let t be a term in normal form. We have $t = C[t_1, \ldots, t_n]$ for some standard terms $t_1, \ldots, t_n$ and an E-context C. The set $Fact_{\mathsf{E}}(t)$ of* factors *of t is defined by $Fact_{\mathsf{E}}(t) = \{t_1, \ldots, t_n\}$. The set $St_{\mathsf{E}}(t)$ of* subterms *of t is the smallest set such that:*

- *$0, t \in St_{\mathsf{E}}(t)$,*
- *if $f(t_1, \ldots, t_n) \in St_{\mathsf{E}}(t)$ is standard then $t_1, \ldots t_n \in St_{\mathsf{E}}(t)$,*
- *if $s \in St_{\mathsf{E}}(t)$ is not standard then $Fact_{\mathsf{E}}(s) \subseteq St_{\mathsf{E}}(t)$.*

Note that the set of factors is uniquely defined since equality is taken to be modulo AC. Note also that, by definition, 0 is not a standard term and the factors of any term are necessarily standard. We extend the notations $St_{\mathsf{E}}(\cdot)$ and $Fact_{\mathsf{E}}(\cdot)$ in a natural way to sets of terms.

Example 1. Let $t_1 = h^2(a)+b+x$ and $t_2 = h(\langle a, b\rangle)+x$, we get $Fact_{\mathsf{E}}(t_1) = \{a, b, x\}$, $St_{\mathsf{E}}(t_1) = \{t_1, a, b, x\}$, $Fact_{\mathsf{E}}(t_2) = \{\langle a, b\rangle, x\}$, $St_{\mathsf{E}}(t_2) = \{t_2, \langle a, b\rangle, a, b, x\}$.

2.3 Proofs

Definition 2. *A* proof *P of $T \vdash u$ is a finite tree such that*

- *the root of P is labeled with $T \vdash u$,*
- *every leaf of P labeled with $T \vdash v$ is such that $v \in T$,*
- *for every node of P labeled with $T \vdash v$ having n sons labeled with $T \vdash v_1, \ldots, T \vdash v_n$, there is an instance $\dfrac{T \vdash v_1 \quad \ldots \quad T \vdash v_n}{T \vdash v}$ (R) of an inference rule. If this node labeled with $T \vdash v$ is the root of P, we say that P* ends *with an instance of* (R).

Note that the terms in the proof are not necessarily ground. A proof P of $T \vdash u$ is *minimal* if there is no proof P' of $T \vdash u$ with less nodes than P.

Definition 3. *A term u is* R-one-step deducible *from a set of terms T in any of the following cases:*

- *$T \vdash u$ is a proof of $T \vdash u$ (i.e, $u \in T$ or $u = 0$),*
- *there exists $u_1, \ldots, u_n$ such that $\dfrac{T \vdash u_1 \quad \ldots \quad T \vdash u_n}{T \vdash u}$ (R) is a proof of $T \vdash u$.*

The term u is one-step deducible *from T if u is R-one-step deducible from T for some inference rule R.*

The following lemma, due to [6], shows that if there exists a proof of a sequent then there exists a "small" one.

Lemma 1. *A minimal proof P of $T \vdash u$ contains only terms in $St_{\mathsf{E}}(T \cup \{u\})$.*

3 Constraint Systems

3.1 Well-Defined Constraint Systems

It is well-known that the security problem of a protocol for a *fixed* number of parallel sessions reduces to the satisfiability of a constraint system (see, e.g. [1, 15]):

Definition 4. *A* constraint *(resp. one-step constraint,* $\mathsf{M_E}$ *constraint) is a sequent of the form* $T \Vdash u$ *(resp.* $T \Vdash_1 u$*,* $T \Vdash_{\mathsf{M_E}} u$*) where* T *is a finite subset of* $\mathcal{T}(\mathcal{F}, \mathcal{X})$ *and* $u \in \mathcal{T}(\mathcal{F}, \mathcal{X})$*. We call* T *the* hypothesis set *of the constraint. A* system of constraints *is a sequence of constraints. A solution to a system* $\mathcal{C}$ *of constraints is a substitution* σ *such that:*

- *for every* $T \Vdash u \in \mathcal{C}$ *there exists a proof of* $T\sigma \vdash u\sigma$*;*
- *for every* $T \Vdash_1 u \in \mathcal{C}$ *the term* $u\sigma$ *is one-step deducible from* $T\sigma$*;*
- *for every* $T \Vdash_{\mathsf{M_E}} u \in \mathcal{C}$ *the term* $u\sigma$ *is* $\mathsf{M_E}$*-one-step deducible from* $T\sigma$*.*

A solution σ to $\mathcal{C}$ is *non-collapsing* if for all $u, v \in St_{\mathsf{E}}(\mathcal{C}) \setminus \mathcal{X}$ such that $u\sigma =_{\mathsf{E}} v\sigma$ then $u =_{\mathsf{E}} v$. If $\mathcal{F}'$ is a sub-signature of $\mathcal{F}$ then a solution σ to a constraint system is called a $\mathcal{F}'$-solution if $x\sigma \in \mathcal{T}(\mathcal{F}', \mathcal{X})$ for every $x \in dom(\sigma)$.

Note that, if σ is solution to a constraint $T \Vdash u$ (resp. one-step constraint, $\mathsf{M_E}$ constraint), then $\sigma\theta$ is also a solution to $T \Vdash u$ for every substitution θ.

Definition 5. *A constraint system* $\mathcal{C} = \{T_i \Vdash u_i\}_{1 \le i \le k}$ *is* well-defined *if:*

1. (monotonicity) *for all* $i < k$*:* $T_i \subseteq T_{i+1}$*,*
2. (origination) *for all substitution* θ*:* $\mathcal{C}\theta$ *satisfies the following requirement:*
$$\forall i \le k, \forall x \in vars(T_i\theta), \exists j < i \text{ such that } x \in vars(u_j\theta).$$

This notion of well-definedness, due to Millen and Shmatikov, is defined in an analogous way on systems of one-step (resp. $\mathsf{M_E}$) constraints. In [16] they show that "reasonable" protocols, in which legitimate protocol participants only execute deterministic steps (up to the generation of random nonces) always lead to a well-defined constraint system. This notion is crucial for several steps of our algorithm.

Theorem 1. *The problem of deciding whether a well-defined constraint system has a solution in* $\mathcal{I}_{\mathsf{DY+E}}$*, where* $\mathsf{E} = \mathsf{ACUNh}$*, is decidable.*

The remainder of the paper is devoted to the proof of this result.

3.2 Conservative Solutions

Intuitively, a *conservative solution* to a constraint system is a solution which does not introduce any new structure. Lemma 2 states that it is sufficient to search for conservative solutions of a constraint system. Moreover, conservative solutions allow us to lift Lemma 1 to deducibility constraints (Lemma 3).

Definition 6. *Let* $\mathcal{C}$ *be a constraint system and* σ *a substitution,* σ *is* conservative *w.r.t.* $\mathcal{C}$ *if and only if for all* $x \in vars(\mathcal{C})$*,* $Fact_{\mathsf{E}}(x\sigma) \subseteq (St_{\mathsf{E}}(\mathcal{C}) \setminus vars(\mathcal{C}))\sigma$*.*

Lemma 2. *Let $\mathcal{C}$ be a well-defined constraint system. If there exists a solution σ to $\mathcal{C}$ then there exists a conservative one.*

Example 2. Consider the following well-defined constraint system $\mathcal{C}$ which is made up of two deducibility constraints: $a, h(b) \Vdash h(x)$ and $a, h(b), x \Vdash \langle a, b \rangle$. One solution is $\sigma = \{x \mapsto \langle a, a \rangle + b\}$. This solution is not conservative w.r.t. $\mathcal{C}$ since $Fact_{\mathsf{E}}(\langle a, a\rangle + b) = \{\langle a, a\rangle, b\}$, and $\langle a, a\rangle$ does not belong to $(St_{\mathsf{E}}(\mathcal{C}) \backslash \{x\})\sigma$. However, as it is said in Lemma 2, there is a conservative solution: $\{x \mapsto b\}$.

Lemma 3. *Let σ be a conservative solution to $\mathcal{C} = \{C_1, \dots, C_k\}$. For each $i \leq k$ there exists a proof of $C_i\sigma$ that involves only terms in $St_{\mathsf{E}}(\mathcal{C})\sigma$.*

4 From Constraints to M_{E} Constraints

We proceed in two non-deterministic steps to reduce the satisfiability of a constraint system to the satisfiability of a M_{E} constraint system:

1. From constraints to one-step constraints (see Lemma 4 and Figure 2).
2. From one-step constraints to M_{E} constraints (see Lemma 5).

```
Input: 𝒞 = {T_1 ⊩ u_1, ..., T_k ⊩ u_k}
  guess S ⊆ St_E(𝒞)
  for all s ∈ S, guess j(s) ∈ {1, ..., k}
  𝒞' := ∅
  for i = 1 to k do
     let S_i := {s | j(s) = i}
     choose a total ordering on S_i (S_i = {s_i^1, ..., s_i^{ki}})
     for j = 1 to ki do
        T := T_i ∪ S_1 ... ∪ S_{i-1} ∪ {s_i^1, ..., s_i^{j-1}}
        𝒞' := 𝒞' ∪ {T ⊩_1 s_i^j}
     end
     𝒞' := 𝒞' ∪ {T ⊩_1 u_i}
  end
return 𝒞'
```

Fig. 2. Step 1: from constraints to one-step constraints.

The idea of the first step is to guess among the subterms of $\mathcal{C}$ those that are going to be deduced by the intruder, and to insert each of them in some order into the constraint system. The completeness of this reduction step is essentially due to the existence of a conservative solution (Lemma 2) and to Lemma 3. In the resulting constraint system, every constraint can be solved by application of a single inference rule:

Lemma 4. *Let $\mathcal{C}$ be a well-defined system of constraints. Let $\mathscr{C}'$ be the set of constraint systems obtained by applying on $\mathcal{C}$ the algorithm described in Figure 2.*

1. *$\mathscr{C}'$ is a finite set of well-defined systems of one-step constraints.*
2. *If some $\mathcal{C}' \in \mathscr{C}'$ has a solution then $\mathcal{C}$ has a solution.*
3. *If $\mathcal{C}$ has a conservative solution then some $\mathcal{C}' \in \mathscr{C}'$ has a conservative solution.*

Lemma 5 allows us to reduce the satisfiability of a system of one-step constraints to the satisfiability of a system of $\mathsf{M_E}$ constraints. We first guess a set R of equalities between subterms. Then, we choose an E-unifier of R among the finite number of possibilities given by Theorem 2.

Theorem 2. *Unification in the theory* ACUNh *is finitary, and there exists an algorithm to compute a complete finite set $mgu_{\mathsf{E}}(R)$ of unifiers of any unification problem R.*

We write $T \vdash_{\mathsf{DY}} u$ if u is (R)-one step deducible from T where R is one of $(\mathsf{D}, \mathsf{UL}, \mathsf{UR}, \mathsf{C})$. It is trivial to decide whether $T \vdash_{\mathsf{DY}} u$ or not. We can now eliminate all constraints $T \Vdash_1 u$ for which $T \vdash_{\mathsf{DY}} u$ already holds.

Lemma 5. *Let $\mathcal{C}$ be a well-defined system of one-step constraints. Let*

$$\mathcal{P} = \{\bigwedge_{(s_1,s_2)\in S'} s_1 = s_2 \mid S' \subseteq St_{\mathsf{E}}(\mathcal{C})^2\}.$$

Let $R \in \mathcal{P}$ and $\theta \in mgu_{\mathsf{E}}(R)$. Let $\mathcal{C}_\theta = \{T\theta \Vdash_{\mathsf{M_E}} u\theta \mid T \Vdash_1 u \in \mathcal{C}$ and $T\theta \not\vdash_{\mathsf{DY}} u\theta\}$. Let $\mathscr{C}$ be the set of constraint systems C_θ obtained this way.

1. *$\mathscr{C}$ is a finite set of well-defined systems of* $\mathsf{M_E}$ *constraints.*
2. *If some $\mathcal{C}_\theta \in \mathscr{C}$ has a solution then $\mathcal{C}$ has a solution.*
3. *If $\mathcal{C}$ has a conservative solution then some $\mathcal{C}_\theta \in \mathscr{C}$ has a non-collapsing solution.*

Note that we can now restrict our attention to *non-collapsing* solutions, thanks to the fact that we have guessed the subterms that are identified by the solution.

5 Solving $\mathsf{M_E}$ Constraints

Now, we have to solve well-defined $\mathsf{M_E}$ constraint systems, where it is sufficient to look for non-collapsing solutions. In the remainder, we consider a $\mathsf{M_E}$ constraint system $\mathcal{C} = \{T_1 \Vdash_{\mathsf{M_E}} u_1, \ldots, T_i \Vdash_{\mathsf{M_E}} u_k\}$ and we assume w.l.o.g. that the set of terms T_i is equal to $\{t_1, \ldots, t_{n+i-1}\}$.

A constraint system is called *factor-preserving* if all its factors appear for the first time in an hypotheses set of a constraint. More formally,

Definition 7. *A* $\mathsf{M_E}$ *constraint system is* factor-preserving *if for all i, $1 \le i \le k$, we have that $Fact_{\mathsf{E}}(u_i) \setminus \mathcal{X} \subseteq \bigcup_{j=1}^{j=n+i-1} Fact_{\mathsf{E}}(t_j)$.*

Example 3. The systems, $\langle a, b\rangle \Vdash_{\mathsf{M_E}} \langle x_1, x_2\rangle$ and $\langle\langle a, b\rangle, a\rangle \Vdash_{\mathsf{M_E}} \langle a, b\rangle$ are not factor-preserving. Note that the first one has no non-collapsing solution whereas the second one has no solution using the $\mathsf{M_E}$ inference rule only.

This notion is important to ensure that well-definedness is maintained when we abstract a constraint system by replacing factors by new constants (see Lemma 7). Fortunately, requiring factor preservation is not a restriction, since:

Lemma 6. *If a well-defined* $\mathsf{M_E}$*-constraint system* $\mathcal{C}$ *has a non-collapsing solution then it is factor-preserving.*

Factor preservation is of course trivial to check. We can hence suppose that the constraint system under consideration is factor-preserving, since if it is not then we conclude immediately by Lemma 6 that it has no non-collapsing solution.

5.1 Reducing the Signature

We will show in Lemma 7 that we can reduce the satisfiability of $\mathsf{M_E}$ constraint systems to the satisfiability of $\mathsf{M_E}$ constraint systems over a signature consisting only of 0, $+$, h, and a set of constants.

If $\rho : M \to N$ is a replacement, that is a bijection between two finite sets of terms M and N, then we denote for any term t by t^ρ the term obtained by replacing in t any top-most occurrence of a subterm $s \in M$ by $s\rho$. This extends in a natural way to constraint systems, and to substitutions.

Lemma 7. *Let* $\mathcal{C}$ *be a well-defined factor-preserving* $\mathsf{M_E}$ *constraint system and* $F = Fact_{\mathsf{E}}(\mathcal{C}) \setminus \mathcal{X}$. *Let* $\mathcal{F}_0$ *be a set of new constant symbols of the same cardinality as* F *and* $\rho : F \to \mathcal{F}_0$ *a bijection.*

1. $\mathcal{C}^\rho$ *is well-defined.*
2. $vars(\mathcal{C}^\rho) = vars(\mathcal{C})$.
3. *If* $\mathcal{C}$ *has a non-collapsing solution then* $\mathcal{C}^\rho$ *has a* $\mathcal{F}_0 \cup \{0, h, +\}$*-solution.*
4. *If* $\mathcal{C}^\rho$ *has a* $\mathcal{F}_0 \cup \{0, h, +\}$*-solution then* $\mathcal{C}$ *has a solution.*

As shown by the example below, well-definedness is not necessarily preserved under abstraction when the system is not factor-preserving.

Example 4. Abstraction of the system $a \Vdash_{\mathsf{M_E}} \langle x_1, x_2 \rangle;\ a, x_1, x_2 \Vdash_{\mathsf{M_E}} b$, which is not factor preserving, yields $a \Vdash_{\mathsf{M_E}} c_{new};\ a, x_1, x_2 \Vdash_{\mathsf{M_E}} b$, which is not well-defined.

5.2 Another Characterization of Well-Definedness

Let $\sum_{i=0}^{n} b_i h^i$ where $b_i \in \mathbb{Z}/2\mathbb{Z}$ be a polynomial of $\mathbb{Z}/2\mathbb{Z}[h]$. The product $\odot$ of a polynomial by a term is a term defined as follows:

$$(\sum_{i=0}^{n} b_i h^i) \odot t = \sum_{i=0 \ \mid\ b_i \neq 0}^{n} h^i(t)$$

For instance $(h^2+1) \odot (x+a) = h^2(x)+x+h^2(a)+a$. Every $t \in \mathcal{T}(\mathcal{F}, \{x_1, \ldots, x_p\})$ can be written $t^{x_1} \odot x_1 + \ldots t^{x_p} \odot x_p + t^0$ with t^{x_v} in $\mathbb{Z}/2\mathbb{Z}[h]$ and $Fact_{\mathsf{E}}(t^0) \cap \mathcal{X} = \emptyset$. We will denote with $\boldsymbol{t}$ the vector $(t^{x_1}, \ldots, t^{x_p})$.

Definition 8. *Let $\mathcal{V} = \{v_1, \ldots, v_m\}$ be a subset of $\mathbb{Z}/2\mathbb{Z}[h]^n$. $\mathcal{V}$ is* independent *if whenever there exist $\alpha_i \in \mathbb{Z}/2\mathbb{Z}[h]$ such that $\alpha_1 v_1 + \ldots + \alpha_m v_m = 0$ then $\alpha_i = 0$ for all $1 \leq i \leq m$. Otherwise $\mathcal{V}$ is* dependent.

Remember that we consider a constraint system $\mathcal{C} = \{t_1, \ldots, t_{n+i-1} \Vdash_{\mathsf{M_E}} u_i\}_{i=1,\ldots,k}$. The set $L = L_k$ of indexes of the so-called *defining constraints* is defined as follow. We set $L_0 = \emptyset$, and we define $L_{i+1} = L_i \cup \{i+1\}$ if $\{\boldsymbol{u_{i+1}}\} \cup \{\boldsymbol{u_j} \mid j \in L_i\}$ is independent, and $L_{i+1} = L_i$ otherwise. We note $\mathcal{B}_i = \{\boldsymbol{u_j} \mid j \in L, j \leq i\}$ and $\mathcal{B} = \mathcal{B}_k$. Lemma 8 gives an algebraic characterization of well-definedness in the special case of the signature $\mathcal{F}_0 \cup \{0, h, +\}$. Now, we have reduced the problem to this restricted signature (Lemma 7), we are going to use the following characterization in Section 5.3 to solve systems of equations over $\mathbb{Z}/2\mathbb{Z}[h]$.

Lemma 8. *A factor-preserving $\mathsf{M_E}$ constraint system $\{t_1, \ldots, t_{n+i-1} \Vdash_{\mathsf{M_E}} u_i\}_{i=1,\ldots,k}$ over the signature $\{0, h, +\} \cup \mathcal{F}_0$ is well-defined if, and only if, for every $i \leq k$, the set of vectors $\{\boldsymbol{t_{n+i-1}}\} \cup \{\boldsymbol{u}_j \mid j \in L_i\}$ is dependent.*

Intuitively, this is related to the fact that matching modulo ACUNh is essentially linear equation solving.

5.3 Solving $\mathsf{M_E}$ Constraint Systems over $\{0, h, +\} \cup \mathcal{F}_0$

We may by Lemma 6 assume that we have a factor-preserving $\mathsf{M_E}$ constraint system. By Lemma 7 satisfiability of such a system can be reduced to satisfiability of a $\mathsf{M_E}$ constraint system over a signature $\{0, h, +\} \cup \mathcal{F}_0$ where $\mathcal{F}_0$ is a finite set of constants. The characterization of Lemma 8 allows us to use the following well-known fact.

Fact 1 *Let A be a matrix $n \times m$ over $\mathbb{Z}/2\mathbb{Z}[h]$ such that the n row vectors are independent ($n \leq m$) then there exists $Q \in \mathbb{Z}/2\mathbb{Z}[h]$ such that*

$$\forall b \in \mathbb{Z}/2\mathbb{Z}[h]^n, \exists X \in \mathbb{Z}/2\mathbb{Z}[h]^m \quad A \cdot X = Q \cdot b \tag{1}$$

Moreover, such a coefficient Q is computable as a determinant of a submatrix of A.

We denote Q_{max} the coefficient Q which satisfies the equation (1) for the matrix $\mathcal{B}$.

Example 5. (running example) To illustrate our procedure, we consider the following well-defined $\mathsf{M_E}$ constraint system:

$$\begin{array}{ll}
h(a)+a, b+h^2(a) & \Vdash_{\mathsf{M_E}} h(x_1)+h^2(x_2) \\
h(a)+a, b+h^2(a), x_1+h(x_2) & \Vdash_{\mathsf{M_E}} x_1+a \\
h(a)+a, b+h^2(a), x_1+h(x_2), h(x_1)+h(a) & \Vdash_{\mathsf{M_E}} h(x_1)+h^2(x_2)+x_1+a
\end{array}$$

We have $\boldsymbol{u_1} = (h, h^2)$, $\boldsymbol{u_2} = (1, 0)$ and $\boldsymbol{u_3} = (1+h, h^2)$. The algorithm returns $L = \{1, 2\}$ and we obtain $Q_{max} = det(\boldsymbol{u}_1, \boldsymbol{u}_2) = h^2$.

Satisfiability of such an $\mathsf{M_E}$ constraint system $\mathcal{C}$ is equivalent to the satisfiability of the following system $\mathcal{S}$ of equations between terms. The variables $z[i,j]$, called *context variables*, take their value in $\mathbb{Z}/2\mathbb{Z}[h]$. Let $\mathcal{Z} = \{z[i,j] \mid 1 \leq i \leq k, 1 \leq j \leq n+i-1\}$.

$$
\begin{aligned}
& z[1,1] \odot t_1 + \ldots + z[1,n] \odot t_n = u_1 \\
& z[2,1] \odot t_1 + \ldots + z[2,n] \odot t_n + z[2,n+1] \odot t_{n+1} = u_2 \\
& \vdots \\
& z[p,1] \odot t_1 + \ldots + z[p,n] \odot t_n + \ldots + z[p,n+p-1] \odot t_{n+k-1} = u_k
\end{aligned}
$$

Example 6. (running example) Let $t_1 = h(a) + a$ and $t_2 = b + h^2(a)$.

$$
\begin{aligned}
& z[1,1] \odot t_1 + z[1,2] \odot t_2 = h(x_1) + h^2(x_2) \\
& z[2,1] \odot t_1 + z[2,2] \odot t_2 + z[2,3] \odot (x_1 + h(x_2)) = x_1 + a \\
& z[3,1] \odot t_1 + z[3,2] \odot t_2 + z[3,3] \odot (x_1 + h(x_2)) + z[3,4] \odot (h(x_1) + h(a)) \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = h(x_1) + h^2(x_2) + x_1 + a
\end{aligned}
$$

Definition 9. *Let $\mathcal{C}$ be a well-defined $\mathsf{M_E}$ constraint system over the signature $\{0, h, +\} \cup \mathcal{F}_0$ and $\mathcal{S}(\mathcal{C})$ be the system of equations obtained from $\mathcal{C}$. A solution to $\mathcal{S}(\mathcal{C})$ is a couple $(\rho : \mathcal{Z} \mapsto \mathbb{Z}/2\mathbb{Z}[h], \theta : vars(\mathcal{C}) \mapsto \mathcal{T}(\{0, h, +\} \cup \mathcal{F}_0))$ such that all the equations of $\mathcal{S}(\mathcal{C})\rho\theta$ are satisfied.*

We split the *context variables* $\mathcal{Z}$ into two parts, those which stem from L and the others. More formally, $\mathcal{Z}_L = \{z[i,j] \mid i \in L \text{ and } 1 \leq j < n+i\}$.

A polynomial $P = \sum_{i=0}^{i=n} p_i h^i$ ($p_n \neq 0$) is *smaller* than $Q = \sum_{i=0}^{i=m} q_i h^i$ ($q_m \neq 0$), written $P < Q$, if either $n < m$, or $P \neq Q$, $n = m$ and $p_i < q_i$ for the greatest i with $p_i \neq q_i$.

Fact 2 *Given any polynomial $P \in \mathbb{Z}/2\mathbb{Z}[h]$, there is only a finite number of polynomials which are smaller (w.r.t. $<$) than P.*

The following Lemma is the crucial point in the proof of Lemma 10.

Lemma 9. *Let $\mathcal{S}(\mathcal{C})$ be a system of equations obtained from a well-defined $\mathsf{M_E}$ constraint system $\mathcal{C}$ over the signature $\{0, h, +\} \cup \mathcal{F}_0$. If $\mathcal{S}(\mathcal{C})$ has a solution then there exists σ a solution to $\mathcal{S}(\mathcal{C})$ such that for all $z \in \mathcal{Z}_L$, $0 \leq z\sigma < Q_{max}$.*

The proof of this lemma proceeds by induction on the number of variables in Z_L.

Lemma 10. *Given $\mathcal{C}$ a well-defined $\mathsf{M_E}$constraint system. It is decidable whether $\mathcal{S}(\mathcal{C})$ has a solution.*

Example 7. (running example) Thanks to Lemma 9, we know that $z[1,1]$, $z[1,2]$, $z[2,1]$, $z[2,2]$ and $z[2,3]$ are bounded by h^2, the value of Q_{max}. We choose $\rho_1 = \{z[1,1] \mapsto 0; z[1,2] \mapsto h; z[2,1] \mapsto h+1; z[2,2] \mapsto 1; z[2,3] \mapsto 0\}$. We do the replacement on the two first equations:

$$
\begin{aligned}
& h \odot (b + h^2(a)) = h(x_1) + h^2(x_2) \\
& (h+1) \odot (h(a) + a) + 1 \odot (b + h^2(a)) = x_1 + a
\end{aligned}
$$

This completely determines the value of x_1 and x_2: $\theta = \{x_1 \mapsto b, x_2 \mapsto h(a)\}$. Lastly, we can apply the substitution θ on the third equation to obtain:

$$z[3,1] \odot (h(a)+a) + z[3,2] \odot (b+h^2(a)) + z[3,3] \odot (b+h^2(a)) + \\ z[3,4] \odot (h(b)+h(a)) = h(b)+h^3(a)+b+a$$

Since this system is linear it is easy to decide whether it has solution.

Let $\rho_2 = \{z[3,1] \mapsto h+1; z[3,2] \mapsto h+1; z[3,3] \mapsto 0; z[3,4] \mapsto 0\}$. The couple $(\rho_1 \cup \rho_2, \theta)$ is a solution to the system of equations described in Example 6.

Now, we are able to prove our main result as stated in Section 3.

Theorem 1. *The problem of deciding whether a well-defined constraint system has a solution in $\mathcal{I}_{\mathsf{DY+E}}$, where $\mathsf{E} = \mathsf{ACUNh}$, is decidable.*

Proof. The procedure described along the paper is sound and complete.

Soundness. Let $\mathcal{C}_1$ be some factor-preserving $\mathsf{M_E}$-constraint system obtained by applying the first part of our procedure on $\mathcal{C}$, a well-defined constraint system. Thanks to Lemma 4 and 5, $\mathcal{C}_1$ is well-defined since $\mathcal{C}$ is well-defined. Let $\mathcal{C}_2$ be the constraint system obtained from $\mathcal{C}_1$ by replacing all factors by different constants. $\mathcal{C}_2$ is well-defined thanks to Lemma 7. Assume that $\mathcal{S}(\mathcal{C}_2)$ (the system of equations associated to $\mathcal{C}_2$) has a solution. We easily deduce that $\mathcal{C}_2$ has a solution, hence by Lemma 7 that $\mathcal{C}_1$ has a solution, and by Lemma 4 and 5 that $\mathcal{C}$ has a solution.

Completeness. Assume that σ is a solution to $\mathcal{C}$. Thanks to Lemma 2, we can assume that σ is conservative w.r.t. $\mathcal{C}$. Let $\mathscr{C}'$ be the finite set of well-defined one-step constraint systems obtained by applying the algorithm described in Section 4 on $\mathcal{C}$. By Lemma 4, we know that there exists $\mathcal{C}' \in \mathscr{C}'$ such that σ is a conservative solution of $\mathcal{C}'$. By Lemma 5, we know that there exists $\mathcal{C}_\theta$ a well-defined $\mathsf{M_E}$-constraint system which has a non-collapsing solution. Hence, $\mathcal{C}_\theta$ is factor-preserving due to Lemma 6. By Lemma 7, $\mathcal{C}_\theta^\rho$ has solution over $\{0, h, +\} \cup \mathcal{F}_0$. Then, Lemma 10 allows us to conclude. □

6 Conclusion

Our solution for solving deducibility constraints is general enough to hold in related equational theories since it relies on general algebraic concepts. In particular, our technique generalizes previous results for the case of the *exclusive or* equational theory ACUN [1, 3] (context variables take values in $\mathbb{Z}/2\mathbb{Z}$) and the theory of Abelian groups AG [16] (contexts are in $\mathbb{Z}$). However, our technique does not apply to the case AGh of the extension of Abelian groups with a homomorphism since then the contexts are in $\mathbb{Z}[h]$, and Fact 2 does not hold. In fact it has recently been shown that this case is undecidable [7].

Despite a superficial similarity between our algorithm and the one of [16], our procedure to reduce $\mathsf{M_E}$-constraints (*cf.* Section 5) to a special class of quadratic equations is different. In particular it makes use of our novel algebraic characterization of well-defined constraint systems. Furthermore, our procedure to solve a particular form of quadratic equations in polynomials over the finite field $\mathbb{Z}/2\mathbb{Z}[h]$ is different from the one proposed in [16].

An open question is the case of an encryption algorithm distributing over *exclusive or*. Although the case of a passive intruder is decidable in this framework [14], the case of an active intruder seems quite intricate since it amounts to having an infinite number of distinct homomorphisms (one for each term used as a key).

References

1. Y. Chevalier, R. Küsters, M. Rusinowitch, and M. Turuani. An NP decision procedure for protocol insecurity with XOR. In *Proc. of 18th Annual IEEE Symposium on Logic in Computer Science (LICS'03)*, pages 261–270. IEEE Comp. Soc. Press, 2003.
2. Y. Chevalier and M. Rusinowitch. Combining intruder theories. In *Proc. 32nd International Colloquium on Automata, Languages and Programming (ICALP'05)*, volume 3580 of *LNCS*, pages 639–651. Springer, 2005.
3. H. Comon-Lundh and V. Shmatikov. Intruder deductions, constraint solving and insecurity decision in presence of exclusive or. In *Proc. of 18th Annual IEEE Symposium on Logic in Computer Science (LICS'03)*, pages 271–280. IEEE Comp. Soc. Press, 2003.
4. V. Cortier, S. Delaune, and P. Lafourcade. A survey of algebraic properties used in cryptographic protocols. *Journal of Computer Security*, 14(1):1–43, 2006.
5. R. Cramer, R. Gennaro, and B. Schoenmakers. A secure and optimally efficient multi-authority election scheme. In *Proc. International Conference on the Theory and Application of Cryptographic Techniques*, volume 1233 of *LNCS*, pages 103–118, 1997.
6. S. Delaune. Easy intruder deduction problems with homomorphisms. *Information Processing Letters*, 97(6):213–218, 2006.
7. S. Delaune. An undecidability result for AGh. Research Report LSV-06-02, LSV, ENS Cachan, France, 2006.
8. S. Delaune and F. Jacquemard. A decision procedure for the verification of security protocols with explicit destructors. In *Proc. of 11th ACM Conference on Computer and Communications Security (CCS'04)*, pages 278–287. ACM Press, 2004.
9. S. Delaune, P. Lafourcade, D. Lugiez, and R. Treinen. Symbolic protocol analysis in presence of a homomorphism operator and *exclusive or*. Research Report LSV-05-20, Cachan, 2005.
10. D. Dolev and A. Yao. On the security of public key protocols. In *Proc. of the 22nd Symp. on Foundations of Computer Science*, pages 350–357. IEEE Comp. Soc. Press, 1981.
11. N. Durgin, P. Lincoln, J. Mitchell, and A. Scedrov. Undecidability of bounded security protocols. In *Proc. Workshop on formal methods in security protocols*, 1999.
12. Q. Guo, P. Narendran, and D. A. Wolfram. Complexity of nilpotent unification and matching problems. *Information and Computation*, 162(1-2):3–23, 2000.
13. P. Lafourcade, D. Lugiez, and R. Treinen. Intruder deduction for AC-like equational theories with homomorphisms. In *Proc. of 16th Int. Conference on Rewriting Techniques and Applications (RTA'05)*, volume 3467 of *LNCS*, pages 308–322. Springer, 2005.
14. P. Lafourcade, D. Lugiez, and R. Treinen. Intruder deduction for the equational theory of exclusive-or with distributive encryption. Research Report LSV-05-19, ENS Cachan, 2005.
15. J. Millen and V. Shmatikov. Constraint solving for bounded-process cryptographic protocol analysis. In *Proc. of 8th ACM Conference on Computer and Communications Security (CCS'01)*. ACM Press, 2001.
16. J. Millen and V. Shmatikov. Symbolic protocol analysis with an Abelian group operator or Diffie-Hellman exponentiation. *Journal of Computer Security*, 13(3):515 – 564, 2005.
17. M. Rusinowitch and M. Turuani. Protocol insecurity with a finite number of sessions, composed keys is NP-complete. *Theoretical Computer Science*, 1-3(299):451–475, 2003.