Deducibility constraints, equational theory and electronic money *

Sergiu Bursuc[1], Hubert Comon-Lundh[1], and Stéphanie Delaune[1,2]

[1] Laboratoire Spécification & Vérification
ENS de Cachan & CNRS UMR 8643, France
{bursuc,comon,delaune}@lsv.ens-cachan.fr
[2] LORIA, CNRS & INRIA project Cassis, Nancy, France

Abstract. The starting point of this work is a case study (from France Télécom) of an electronic purse protocol. The goal was to prove that the protocol is secure or that there is an attack. Modeling the protocol requires algebraic properties of a fragment of arithmetic, typically containing modular exponentiation. The usual equational theories described in papers on security protocols are too weak: the protocol cannot even be executed in these models. We consider here an equational theory which is powerful enough for the protocol to be executed, and for which unification is still decidable.
Our main result is the decidability of the so-called intruder deduction problem, *i.e.* security in presence of a passive attacker, taking the algebraic properties into account. Our equational theory is a combination of several equational theories over non-disjoint signatures.

1 Introduction

The formal verification of security protocols is now a well-established area of research. One of the main challenges during the past years was to refine the models, taking into account some algebraic properties of the cryptographic primitives. Representing messages as terms in a free algebra, which is known as the *perfect cryptography assumption*, allows to find some logical attacks, but fails to find some of them, which rely on the algebraic properties. Even worse, for some protocols, the local program of an honest agent may itself rely on some algebraic properties. In such a case, the protocol does not have any honest execution in the perfect cryptography model.

More precise models of protocols are therefore needed. They assume that messages are terms modulo an equational theory. A list of some relevant equational theories is proposed in [9] (see also [18] for a list of protocols and comments on possible attacks).

Proving the security for a bounded number of sessions in such formal models deserved a lot of articles, which we cannot all cite here. Let us mention [16], in which the authors proved that the security problem is co-NP complete in the

* This work has been partly supported by the RNTL project PROUVÉ and POSÉ.

perfect cryptography case. The extension to several equational theories has been considered: exclusive-or [7, 2], Abelian groups [17], some properties of modular exponentiation [3, 15], homomorphisms and exclusive-or [10],... All these works rely on solving *deducibility constraints* modulo equational theories, an approach that we will follow in the present paper.

On the other side, if we put too much of arithmetic in the equational theory, getting a more precise model, the security problem becomes undecidable: a necessary condition is the decidability of unification. A typical problem is: which properties of modular exponentiation do we want to keep? As shown in [13], the boundary between decidability and undecidability is tight.

We are interested here in yet other properties of modular exponentiation. In a case study of an electronic purse protocol (whose some parts will be described in Section 2) submitted by France Télécom, the protocol cannot be even executed if we don't have both the properties $(x^y)^z = x^{y\times z}$ and $x^y \times x^z = x^{y+z}$, as well as some other properties described later. However, having both multiplication and addition of exponents, together with the usual distributivity laws, yields undecidability of unification by an easy encoding of integer arithmetic. Nevertheless, we managed to design some equational theory for which unification is decidable and the protocol can be executed. The theory will be described in detail in Section 2. It is a union of three Abelian group theories and some rules for exponentiation.

Our equational theory does not fall in any class for which the security problem is known to be decidable. In view of the number of symbols and rules, it is worth trying to use combination results. Unfortunately, we cannot use directly the results of [4], as our theories are not disjoint. Further (closer) results are those of Y. Chevalier and M. Rusinowitch in [5], in which the authors give combination results for non-disjoint signatures, with applications to some security issues in presence of modular exponentiation. However, again, we cannot apply these results, as our theory can not be split into two equational theories satisfying the hypotheses of [5].

We were left to develop a new decision procedure. An important step towards this result is to decide the so-called *intruder deduction problem*: Given a finite set of messages T and a given message m, is it possible for the intruder to retrieve m from T by using his deduction capabilities? This corresponds to the security decision problem in presence of a passive eavesdropper, *i.e.* an intruder who is only able to listen messages that pass over the network. In particular it is assumed that he can not intercept messages and send some fake messages over the network. In this paper, we propose a decision procedure to decide this problem in presence of an intruder having complex deduction capabilities which are modeled through an equational theory. This is achieved by using a locality lemma from which it follows that the intruder deduction problem can be decided in polynomial time.

2 Intruder deduction problem

In this section, we describe our case study and the equational theory allowing us to model the protocol. Then, we formally describe the problem we are interested in. Our main result is stated in Theorem 1.

2.1 The electronic purse protocol

The protocol involves three possible agents: the electronic purse *EP*, a server S and a trusted authority A. We will not consider here the authority A, who is involved only in case of claims of either party (and we also simplify several parts in the following). We denote by b and r two positive integers, which are public. The public key of *EP* is $b^s \mod r$ whereas s is its private key.

First, there is a phase during which the server authenticates itself. We skip this phase here, which does not make use of algebraic properties. After this phase, S and *EP* agree on a session nonce N_s and S owes the (certified) public key $b^s \mod r$ of *EP*. Then

1. The purse *EP* generates a nonce N, builds a message M (which is only used in case of conflict and whose content is not relevant here) and sends to the server S: $\mathsf{hash}(b^N \mathsf{mod}\ r, S, N_s, M, X)$, where X is the amount payed.
2. The server S challenges EP sending a nonce N_c.
3. The purse EP sends back $N - s \times N_c, M, X$ and subtract X from his account.
4. The server S checks that the message received at step 1 is consistent with the message received at step 3 and then increases his account from the amount X.

The important and difficult part is the last step: S should be able to complete this verification. Here are the operations performed by S at this stage:

$$\mathsf{hash}((b^s)^{N_c} \times b^{N-s\times N_c} \mathsf{mod}\ r, S, N_s, M, X) = \mathsf{hash}(b^N \mathsf{mod}\ r, S, N_s, M, X)$$

The server S raises b^{-s} to the power N_c (b^s is public and N_c is known), raises b to the power $N - s \times N_c$ (which is the message sent at step 3), and multiply the two results. We can see that the following equational properties are used:

$$\mathsf{exp}(\mathsf{exp}(b, y), z) = \mathsf{exp}(b, y \times z) \quad \mathsf{exp}(b, x) \times \mathsf{exp}(b, y) = \mathsf{exp}(b, y + z)$$

as well as Abelian group properties of both $\times$ and $+$.

2.2 The equational theory

The problem now is that if we put together the above properties and the Abelian group properties of $+$ and $\times$, we can derive the distributivity of $\times$ w.r.t. $+$, in which case unification (hence security) becomes undecidable (see *e.g.* [9]). That is why we used a first trick: we introduce a unary function symbol h, whose meaning is $h(x) = \mathsf{exp}(b, x)$. We also use two distinct multiplication symbols: $\bullet$ and $\star$, with the following equational axioms EP: $\mathsf{AG}(+, J_+, e_+)$, $\mathsf{AG}(\star, J_\star, e_\star)$, $\mathsf{AG}(\bullet, J_\bullet, e_\bullet)$

(where AG are the axioms of Abelian Groups, which will be discussed later) as well as:

$$\begin{aligned} \mathsf{exp}(h(x), y) &= h(x \star y) \qquad\qquad & h(x) \bullet h(y) = h(x+y) \\ \mathsf{exp}(\mathsf{exp}(x, y), z) &= \mathsf{exp}(x, y \star z) & \end{aligned}$$

These equational axioms suffice for the verification at the last step of the protocol. The distinction of the two multiplication symbols is not necessary for the purpose of the present paper: everything holds if we equate $\bullet$ and $\star$. However, we try here to meet the conditions of [5] for the combination results: the distinction between the two multiplication symbols might be useful when extending the results of this paper to the active intruder case.

It remains to show that unification is decidable modulo this theory. This is the subject of Section 3.

2.3 Security problem

The most widely used deduction relation representing the deduction abilities of an intruder is often referred to as the Dolev-Yao model [12]. However, we want to give to the intruder the power to use equational reasoning modulo the set EP of equational axioms. The resulting set of deduction rules, denoted by $\mathcal{I}_{\mathsf{EP}}$ is given in Figure 1 where $\mathcal{F} = \{+, J_+, \star, J_\star, \bullet, J_\bullet, \mathsf{exp}, h\}$. This is the now classical approach, using explicit destructors. When f is associative and commutative, the number of premises of such a rule is unbounded; the set of intruder deduction rules is recursive (but might be infinite).

$$\frac{T \vdash u_1 \quad \dots \quad T \vdash u_n}{T \vdash f(u_1, \dots, u_n)} \text{ where } f \in \mathcal{F} \qquad\qquad (\mathsf{Eq})\ \frac{T \vdash u}{T \vdash v}\ u =_{\mathsf{EP}} v$$

Fig. 1. Inference system — $\mathcal{I}_{\mathsf{EP}}$

Assume given an intruder theory. The problem whether an intruder can gain certain information s from a set of knowledge T, *i.e.* whether there is a proof of $T \vdash s$, is called the *intruder deduction problem.*

INPUT: a finite set of terms T, a term s (the secret).
OUTPUT: Does there exist a proof of $T \vdash s$?

Theorem 1. *The intruder deduction problem is decidable in polynomial time for the inference system $\mathcal{I}_{\mathsf{EP}}$.*

To prove this result, we will first introduce a new inference system that is equivalent from the point of view of deduction. Indeed, the proof system given in Figure 1 is not appropriate for automated proof search: the rule (Eq) allows equational reasoning at any moment of a proof. To define a more effective model, we represent the equational theory by an AC-convergent rewrite system. The rewriting system together with some of its properties are given in Section 3. Moreover, in order to make easier some reasoning we will split the rule about exp into three different inference rules. This new inference system will be fully described at the beginning of Section 4.

3 Properties of the equational theory

In this section we study the equational theory we have introduced in Section 2.2. We show that this theory can be represented by an AC-convergent rewriting system and we establish that unification modulo EP is decidable. Lastly, we prove some technical lemmas which will be useful to establish our locality result stated in Proposition 2. We rely on classical results on rewriting modulo equations (in particular modulo AC). See [11] for the definitions and notations.

3.1 Rewriting system associated to the equational theory EP

For simplicity, our alphabet will contain a finite number of free constant symbols and the associative-commutative symbols $\{\star, \bullet, +\}$, the binary symbol exp, the unary symbols $h, J_\star, J_+, J_\bullet$ and the 3 neutral elements. We could also add other symbols, such as encryption, hashing,... and use then combination results of [1] allowing us to conclude in the case of disjoint theories.

The equational theory EP can actually be presented by a finite convergent rewrite system $\mathcal{R}$ (modulo associativity and commutativity (AC) of $+$, $\star$ and $\bullet$), which has actually even stronger properties. First, for each $\circ \in \{+, \star, \bullet\}$ $\mathcal{R}_{\mathsf{AG}(\circ)}$ is the rewrite system modulo AC for $\circ$:

$$\begin{array}{cc}
x \circ e_\circ \to x & x \circ J_\circ(x) \to e_\circ \\
J_\circ(x) \circ J_\circ(y) \to J_\circ(x \circ y) & J_\circ(e_\circ) \to e_\circ \\
J_\circ(J_\circ(x)) \to x & J_\circ(x) \circ x \circ y \to y \\
J_\circ(x) \circ J_\circ(y) \circ z \to J_\circ(x \circ y) \circ z & J_\circ(x \circ y) \circ x \to J_\circ(y) \\
J_\circ(x \circ y) \circ x \circ z \to J_\circ(y) \circ z & J_\circ(J_\circ(x) \circ y) \to x \circ J_\circ(y)
\end{array}$$

where $e_\circ$ is the appropriate neutral element. The unusual orientation of rules for inverses will ensure strong properties of the rewrite system, as explained in [6]. In addition, we have the following rewrite rules:

$$\mathcal{R}_0 = \left\{ \begin{array}{cc}
\mathsf{exp}(h(x), y) \to h(x \star y) & J_\bullet(h(x)) \to h(J_+(x)) \\
\mathsf{exp}(\mathsf{exp}(x, y), z) \to \mathsf{exp}(x, y \star z) & h(e_+) \to e_\bullet \\
h(x) \bullet h(y) \to h(x + y) & J_\bullet(h(x) \bullet y) \to h(J_+(x)) \bullet J_\bullet(y) \\
h(x) \bullet h(y) \bullet z \to h(x + y) \bullet z & \mathsf{exp}(e_\bullet, x) \to h(e_+ \star x)
\end{array} \right.$$

The rewriting system $\mathcal{R} = \mathcal{R}_{\mathsf{AG}(\star)} \cup \mathcal{R}_{\mathsf{AG}(\bullet)} \cup \mathcal{R}_{\mathsf{AG}(+)} \cup \mathcal{R}_0$ consists of the 38 rewrite rules and the following result has been mechanically verified using CiME [8].

Lemma 1. *$\mathcal{R}$ is convergent modulo associativity and commutativity.*

The normal form (modulo AC) of t is written $t{\downarrow}$. Furthermore, not only $\mathcal{R}$ is convergent, but also:

Lemma 2. *$(\mathcal{R}, \mathsf{AC})$ is a decomposition of the equational theory EP which has the finite variant property.*

This property has been introduced in [6] and ensures that, for any term (or finite set of terms) t, there is a finite computable set of substitutions $\theta_1, \ldots, \theta_n$ such that, for any substitution σ, there exists an index i and a substitution σ' such that $t\sigma{\downarrow} = t\theta_i{\downarrow}\sigma'$. In other words, all possible reductions in an instance of t can be computed in advance. The lemma can be proved using a sufficient condition introduced in [6] and called *boundedness*. The interest of this property is twofold. First, due to the fact that unification is decidable for the theory AC, it ensures that unification is also decidable for EP. Secondly, such a property will be certainly useful to lift our result to solve intruder deduction constraints with variables in order to decide the security problem in presence of an active attacker.

3.2 Notion of subterm

We assume the reader familiar with the basic vocabulary and results on term rewriting systems and term rewriting systems modulo AC. As usual, AC symbols are also considered as variadic symbols and may be used in infix notation and terms are flattened. For $\circ \in \{\star, +, \bullet\}$, we define $\mathsf{inv}_\circ(u)$ as the term $J_\circ(u){\downarrow}$. For instance, we have that $\mathsf{inv}_\bullet(h(J_+(a))) = J_\bullet(h(J_+(a))){\downarrow} = J_\bullet(h(J_+(a))){\downarrow} = h(a)$.

Definition 1. *We denote by $top(t)$ the root symbol of the term t. $\textsc{top}(u)$ is defined by $\textsc{top}(J_\circ(v \circ w)) = \circ$, $\textsc{top}(h(w + v)) = \bullet$, $\textsc{top}(h(J_+(u + v))) = \bullet$ and $\textsc{top}(u) = top(u)$ otherwise.*

For instance, we have that $\textsc{top}(h(a{+}b)) = \bullet$, $\textsc{top}(h(a)) = h$, $\textsc{top}(J_+(a{+}b)) = +$ and $\textsc{top}(J_+(a)) = J_+$.

Definition 2. *Let $\circ \in \{\star, +, \bullet\}$, the set $\mathrm{DS}_\circ(u)$ is defined by*

- $\mathrm{DS}_\circ(u \circ v) = \mathrm{DS}_\circ(u) \cup \mathrm{DS}_\circ(v)$,
- $\mathrm{DS}_\circ(J_\circ(u)) = \{J_\circ(v) \mid v \in \mathrm{DS}_\circ(u)\}$,
- $\mathrm{DS}_\bullet(h(u)) = \{h(v) \mid v \in \mathrm{DS}_+(u)\}$, *and*
- $\mathrm{DS}_\circ(u) = \{u\}$ *if* $\textsc{top}(u) \neq \circ$.

In particular, note that $\mathrm{DS}_\bullet(h(J_+(a + b))) = \{h(J_+(a)),\ h(J_+(b))\}$.

Definition 3 (subterms). *Let t be a term in normal form, $Sub(t)$ is the smallest set of terms such that $t \in Sub(t)$ and if $u \in Sub(t)$ then*

- *either* $\circ = \textsc{top}(u) \in \{\star, \bullet, +\}$ *and* $\mathrm{DS}_\circ(u) \subseteq Sub(t)$
- *or else* $u = f(u_1, \ldots, u_n)$ *and* $u_1, \ldots, u_n \in Sub(t)$.

This notion is extended as expected to set of terms.

Example 1. Let $t_1 = J_+(a + b)$, $t_2 = h(J_+(b))$, $t_3 = J_\star(J_+(b)) \star c$ and $t_4 = h(c)$. We have that $\mathrm{Sub}(t_1) = \{t_1,\ J_+(a),\ J_+(b),\ a,\ b\}$, $\mathrm{Sub}(t_2) = \{t_2,\ J_+(b),\ b\}$, $\mathrm{Sub}(t_3) = \{t_3,\ J_\star(J_+(b)),\ J_+(b),\ b,\ c\}$, and $\mathrm{Sub}(t_4) = \{t_4,\ c\}$.

3.3 Technical lemmas on rewriting

The lemmas stated and proved below are used in the proof of Proposition 2.

Lemma 3. *Let $t, t_1, \ldots, t_n$ be terms in normal form, $n \geq 1$, $\circ \in \{\star, \bullet, +\}$, $\textsc{top}(t) \notin \{\circ, e_\circ\}$. Assume that $\textsc{top}((t \circ t_1 \circ \ldots \circ t_n)\downarrow) \neq \circ$ and $(t_1 \circ \ldots \circ t_n)\downarrow \neq e_\circ$. Then, there is an index i such that $\mathsf{inv}_\circ(t) \in \mathrm{DS}_\circ(t_i)$.*

Proof. The rewrite system $\mathcal{R}$ is convergent. So we can choose a strategy for reducing $t \circ t_1 \circ \ldots \circ t_n$ to its normal form. Given a term u, we order possible redexes $l\sigma \to r\sigma$ in u increasing order of priority as follows:

1. $l = J_\circ(x) \circ J_\circ(y) \circ z$ and $J_\circ(x)\sigma = t$ (or $J_\circ(y)\sigma = t$, the t in the lemma's statement)
2. $l = h(x) \bullet h(y) \bullet z$ and $h(x)\sigma = t$ (or $h(y)\sigma = t$)
3. $l = h(x) \bullet h(y) \bullet z$, and $h(x)\sigma \neq t, h(y)\sigma \neq t$
4. all other cases

We contract always a redex with a maximal priority. This means that the first two cases are applied, only when other rules instances are not a redex in u.

Then we prove the result on the length of such a reduction sequence of $t \circ t_1 \circ \ldots \circ t_n$ to its normal form.

The case where the reduction length is 0 does not occur. Now, we investigate the possible rules, which are applied for the first reduction step. There are 7 cases when $\circ \neq \bullet$ and two additional cases when $\circ = \bullet$:

Case 1 : The rule is $x \circ e_\circ \to x$. Since $t, t_1, \ldots, t_n$ are in normal form, we must have $t_i = e_\circ$ for some i. We simply apply the induction hypothesis. Note that, because $(t_1 \circ \ldots \circ t_n)\downarrow \neq e_\circ$, n must be at least 2.

Case 2 : The rule is $x \circ J_\circ(x) \to e_\circ$: $t \circ t_1 \circ \ldots \circ t_n = x\sigma \circ J_\circ(x\sigma)$. Either $t = J_\circ(x\sigma)$ and, since $\textsc{top}(t) \neq \circ$, we must have $n = 1$ and $t_1 = \mathsf{inv}_\circ(t)$ or else there is an index i such that $t_i = J_\circ(x\sigma)$ or $t_i = J_\circ(x\sigma) \circ t'_i$. In the first case, $x\sigma = t \circ u$ and then $t_i = J_\circ(t \circ u)$. In the second case either $x\sigma = t \circ u$ and $t_i = J_\circ(t \circ u) \circ t'_i$ or $t'_i = t \circ u$ and $t_i = t \circ u \circ J_\circ(x\sigma)$.

Case 3 : The rule is $J_\circ(x) \circ x \circ y \to y$. Then, as in case 2, $t = J_\circ(x\sigma)$ and, for some i, $t_i = \mathsf{inv}_\circ(t) \circ u$ or else there is an index i such that $t_i = J_\circ(x\sigma) \circ t'_i$. In that case, either $t_i = t \circ u$ or $y\sigma = t \circ t'_1 \circ \ldots \circ t'_k$ and, for each k, t'_k is in normal form and there is an injection π from $\{1, \ldots, k\}$ in $\{1, \ldots, n\}$ such that, for every j, $t_{\pi(j)} = t'_j \circ u_j$. Moreover, $(t'_1 \circ \ldots \circ t'_k)\downarrow = (t_1 \circ \ldots \circ t_n)\downarrow \neq e_\circ$. Then, we can apply the induction hypothesis: there is an index j such that $t'_j = \mathsf{inv}_\circ(t)$ or $t'_j = \mathsf{inv}_\circ(t) \circ u$ or $t'_j = J_0(t \circ u) \circ v$ or $t'_j = J_0(t \circ u)$. In each case, choosing $i = \pi(j)$, we get the desired properties.

Case 4 : The rule is $J_\circ(x) \circ J_\circ(y) \circ z \to J_\circ(x \circ y) \circ z$. If $t = J_\circ(x)\sigma$, then, by hypothesis on the strategy, $J_\circ(y\sigma) \circ z\sigma$ is in normal form and, moreover, $z\sigma$ cannot be written $t \circ u$ (otherwise another rule applies). Then $J_\circ(x \circ y)\sigma \circ z\sigma$ is in normal form, which contradicts $\textsc{top}((t \circ t_1 \circ \ldots \circ t_n)\downarrow) \neq \circ$: this case cannot occur.

Now, $J_\circ(x \circ y)\sigma$ is in normal form and $z\sigma = t \circ t'_1 \circ \ldots \circ t'_k$ where the terms $t'_1, \ldots t'_k$ are in normal form and $(J_\circ(x \circ y)\sigma \circ t'_1, \ldots \circ t'_k)\downarrow \neq e_\circ$. We can apply the induction hypothesis: $\mathsf{inv}_\circ(t) \in \mathrm{DS}_\circ(t'_j)$ for some j or else $\mathsf{inv}_\circ(t) \in \mathrm{DS}_\circ(J_\circ(x \circ y)\sigma)$. In the first case, as in case 3, there is some index $i = \pi(j)$ such that $t_i = t'_j \circ u$, hence $t_i = \mathsf{inv}_\circ(t) \circ v \circ u$. In the second case, there are indices i_1, i_2 such that $t_{i_1} = J_\circ(x\sigma) \circ u$ and $t_{i_2} = J_\circ(y\sigma) \circ v$ (u and v might be empty here). Hence there is a variable (say x) such that $t = J_\circ(t')$ and $x\sigma = t' \circ u$. Then $t \in \mathrm{DS}_\circ(t_{i_1})$.

Case 5 : The rule is $J_\circ(x) \circ J_\circ(y) \to J_\circ(x \circ y)$. This case cannot occur since the resulting term would be in normal form (remember $J_\circ(x\sigma)$ and $J_\circ(y\sigma)$ are assumed both in normal form) and we would not have $\textsc{top}((t \circ t_1 \circ \ldots \circ t_n)\downarrow) = \circ$.

Case 6 : The rule is $J_\circ(x \circ y) \circ x \to J_\circ(y)$. In this case, t cannot be $J_\circ(x \circ y)\sigma$ since $\textsc{top}(t) \neq \circ$. Hence $x\sigma = t \circ u$ and there is an index i such that $t_i = J_\circ(x \circ y)\sigma \circ v$ (with possibly empty u or v). Then $t \in \mathrm{DS}_\circ(t_i)$.

Case 7 : The rule is $J_\circ(x \circ y) \circ x \circ z \to J_\circ(y) \circ z$. As in case 6, $J_\circ(x \circ y)\sigma$ cannot be t itself: either $x\sigma = t \circ u$ or else $z\sigma = t \circ u$ for some (possibly empty) u. Moreover, there is an index i such that $t_i = J_\circ(x \circ y)\sigma \circ w$ (with possibly empty w).

In the first case, $t \in \mathrm{DS}_\circ(t_i)$ and, in the second case, we apply the induction hypothesis: if $z\sigma = t'_1 \circ \ldots \circ t'_k$, either there is some index j such that $\mathsf{inv}_\circ(t) \in \mathrm{DS}_\circ(t'_j)$, in which case, as before, there is some index k such that $\mathsf{inv}_\circ(t) \in \mathrm{DS}_\circ(t_k)$ or else $\mathsf{inv}_\circ(t) \in \mathrm{DS}_\circ(J_\circ(y\sigma))$. Then $t \in \mathrm{DS}_\circ(t_i)$.

Case 8 : the rule is $h(x) \bullet h(y) \to h(x+y)$: let $h(x)\sigma = t = h(u_1)$, $h(y)\sigma = h(u_2)$ and $\mathrm{top}(u_1) \neq +$. According to the strategy, $h(y)\sigma$ is in normal form.Since $h(e_+) \to e_\bullet$ and $h(u_1 + u_2)\downarrow \neq e_\bullet$, $h(u_1 + u_2)\downarrow = h((u_1 + u_2)\downarrow)$ and we can apply the induction hypothesis to $u_1 + u_2$ (with $\circ = +$): $\mathsf{inv}_+(u_1) \in \mathrm{DS}_+(u_2)$, which implies $h(\mathsf{inv}_+(u_1)) \in \mathrm{DS}_\bullet(h(u_2))$. But $\mathsf{inv}_\bullet(h(u_1)) = h(\mathsf{inv}_+(u_1))$ by definition. Hence $\mathsf{inv}_\bullet(t) \in \mathrm{DS}_\bullet(h(u_2))$, which is the desired result.

Case 9 : the rule is $h(x) \bullet h(y) \bullet z \to h(x + y) \bullet z$. If we had $h(x)\sigma = t$ (resp. $h(y)\sigma = t$), by hypothesis on the strategy, we would have $h(y)\sigma \bullet z\sigma$ in normal form. In particular, $z\sigma$ cannot be written $h(v) \bullet w$ or $h(v)$ or $J_\bullet(h(v)) \bullet w$ or $J_\bullet(h(v))$. This implies that $(t \circ \ldots \circ t_n)\downarrow = h(x\sigma + y\sigma)\downarrow \bullet z\sigma$, contradicting $\mathrm{top}((t \circ t_1 \circ \ldots \circ t_n)\downarrow) \neq \circ$.

It follows that $z\sigma = t \bullet t'_1 \bullet \ldots \bullet t'_k$ (and each t'_i is some t_j). Moreover, each t'_i and t itself must be headed with h (by the assumed strategy and since the normal form is not headed with $\bullet$). Let $t_i = h(u_i)$ and $t = h(u_0)$. $(t \bullet t_1 \ldots \bullet t_n)\downarrow = h(v)$ and $(u_0 + \ldots + u_n)\downarrow = v$. Now, $(u_1 + \ldots + u_n)\downarrow \neq e_+$ and $\textsc{top}((u_0 + \ldots + u_n)\downarrow) \neq +$. Hence, by induction hypothesis, $\mathsf{inv}_+(u_0) \in \mathrm{DS}_+(u_i)$ for some i. It follows that $\mathsf{inv}_\bullet(t) \in \mathrm{DS}_\bullet(t_i)$. $\square$

Lemma 4. *Let* $\circ \in \{\star, \bullet, +\}$, $t_1, \ldots, t_n, u_1, \ldots, u_m$ *be terms in normal form such that for every* i, $\textsc{top}(t_i) = \circ$ *and* $\textsc{top}(u_i) \notin \{\circ, e_\circ\}$. *Let* $t = (t_1 \circ \ldots \circ t_n \circ u_1 \circ \ldots \circ u_m)\downarrow$. *Then for every* i, *either* $u_i \in \mathrm{DS}_\circ(t)$ *or there is an index* j *such that* $\mathsf{inv}_\circ(u_i) \in \mathrm{DS}_\circ(t_j)$, *or else there is an index* j *such that* $u_j = \mathsf{inv}_\circ(u_i)$.

Proof. We use the lemma 3, with u_i in place of t and adding a term $t_{n+1} = \mathsf{inv}_\circ(t)$: we conclude that, for every i, either $\mathsf{inv}_\circ(u_i) \in \mathrm{DS}_\circ(t_{n+1})$ or $\mathsf{inv}_\circ(u_i) \in \mathrm{DS}_\circ(u_j)$ or $\mathsf{inv}_\circ(u_i) \in \mathrm{DS}_\circ(t_j)$. In the first case, $u_i \in \mathrm{DS}_t()$, in the second case $\mathsf{inv}_\circ(u_i) = u_j$. $\square$

4 Locality

We first introduce a new inference system equivalent to $\mathcal{I}_{\mathsf{EP}}$ and then we will show that this inference system is local w.r.t. to a notion of subterms F.

Definition 4 (F-local). *An inference system $\mathcal{I}$ is F-local if for any proof of $T \vdash u$ in $\mathcal{I}$ there exists one such that all intermediate formulas are in $F(T \cup \{u\})$.*

4.1 A local inference system

We introduce a new inference system which can be viewed as the union of two parts denoted respectively by $\mathcal{I}_1$ and $\mathcal{I}_2$. From now on we omit the rule (Eq) and consider a variant of the deduction model which works on normal forms. This means that, after each step, the term obtained is reduced to its normal form. The part $\mathcal{I}_1$ is made up of the following 7 rules where $\mathcal{F}^- = \{+, J_+, \star, J_\star, \bullet, J_\bullet, h\}$.

$$\mathcal{I}_1 = \left\{ (\mathsf{R}_f)\ \frac{T \vdash u_1 \quad \dots \quad T \vdash u_n}{T \vdash f(u_1, \dots, u_n)\downarrow} \text{ where } f \in \mathcal{F}^- \right.$$

We also distinguish the rules obtained by exponentiation, depending on the first premise of the inference rule: either applying exponentiation to u, v yields a term $\mathsf{exp}(u, v)$ in normal form or else $u = \mathsf{exp}(t_1, t_2)$ or else $u = h(t_1)$. We distinguish these three cases splitting the single inference rule into three different inference rules, which will be more convenient for further proofs. We let $\mathcal{I}_2$ be the inference system made up of the three following rules:

$$\mathcal{I}_2 = \left\{ \frac{h(t_1)\ \ t_2\ \ \cdots\ \ t_n}{h(t_1 \star \cdots \star t_n)\downarrow}\,\mathsf{Exp}_1 \qquad \frac{\mathsf{exp}(t_1, t_2)\ \ t_3\ \ \cdots\ \ t_n}{\mathsf{exp}(t_1, t_2 \star \cdots \star t_n)\downarrow}\,\mathsf{Exp}_2 \qquad \frac{t\ \ u}{\mathsf{exp}(t, u)}\,\mathsf{Exp}_3 \right.$$

Equivalence modulo AC is easy to decide, so we omit the equality rule for AC and just work with equivalence classes modulo AC. We have the following result.

Proposition 1. *Let T be a set of terms and u a term (in normal form). We have that $T \vdash u$ is derivable in $\mathcal{I}_{\mathsf{EP}}$ if and only if $T \vdash u$ is derivable in $\mathcal{I}_1 \cup \mathcal{I}_2$.*

Definition 5 (decomposition rule). *The application of a rule in $\mathcal{I}_2$ is a* decomposition *if it is an instance of Exp_1 and the resulting term is of the form $h(u)$ with $\textsc{top}(u) \neq \star$. A* decomposition rule *for $\mathcal{I}_1$ is a rule R_f, such that one of the following occurs:*

- *$f \in \{\star, \bullet, +\}$ and the conclusion $t = (f(t_1, \dots, t_n))\downarrow$ is such that $\textsc{top}(t) \neq f$*
- *$f = J_\circ$ and the rule is applied to a term of the form $J_\circ(t)$*

Rules, which are not decomposition rules are compositions.

4.2 Locality result

We show that our case study satisfies the locality properties. First we need to define a suitable function F. We consider the following one:

$$
\begin{aligned}
F(T) = \quad & \mathrm{Sub}(T) \\
& \cup \{h(t) \mid t \in \mathrm{Sub}(T), \textsc{top}(t) = +\} \\
& \cup \{h(\mathsf{inv}_+(t) \mid t \in \mathrm{Sub}(t), \textsc{top}(t) = +\} \\
& \cup \{\mathsf{inv}_\circ(t) \mid t \in \mathrm{Sub}(T), \textsc{top}(t) = \circ, \circ \in \{\star, +\}\} \\
& \cup \{h(t) \mid \exists u \in \mathrm{Sub}(T) \text{ such that } \textsc{top}(u) = \circ, t \in \mathrm{DS}_\circ(u), \circ \in \{\star, +\}\} \\
& \cup \{\mathsf{inv}_\circ(t) \mid \exists u \in \mathrm{Sub}(T) \text{ such that } \textsc{top}(u) = \circ, t \in \mathrm{DS}_\circ(u), \circ \in \{\star, +, \bullet\}\} \\
& \cup \{h(\mathsf{inv}_\circ(t)) \mid \exists u \in \mathrm{Sub}(T) \text{ such that } \textsc{top}(u) = \circ, t \in \mathrm{DS}_\circ(u), \circ \in \{\star, +\}\}
\end{aligned}
$$

Lemma 5. *The size of $F(T)$ (number of distinct subterms) is linear in the size of T.*

Proof. More precisely, the size of $F(T)$ is bounded by 10 times the size of T. For, it suffices to note that, all terms in $F(T)$ are always in $\mathrm{Sub}(T) \cup h(\mathrm{Sub}(T)) \cup \mathsf{inv}_\circ(\mathrm{Sub}(T)) \cup h(\mathsf{inv}_\circ(\mathrm{Sub}(T)))$ for some $\circ$.

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

Proposition 2. *The inference system $\mathcal{I}_1 \cup \mathcal{I}_2$ is F-local.*

Example 2. Here are some examples of proofs, which satisfy the requirements of the proposition:

$$
\dfrac{h(a+b+c) \quad \dfrac{\dfrac{a}{J_+(a)}\,\mathsf{R}_{J_+}}{h(J_+(a))}\,\mathsf{R}_h}{h(b+c)}\,\mathsf{R}_\bullet
\qquad
\dfrac{\dfrac{a+b+c}{h(a+b+c)}\,\mathsf{R}_h \quad h(J_+(a))}{h(b+c)}\,\mathsf{R}_\bullet
$$

$$
\dfrac{h(b) \quad \dfrac{a \star b}{J_\star(a \star b)}\,\mathsf{R}_{J_\star}}{h(J_\star(a))}\,\mathsf{Exp}_1
\qquad
\dfrac{\dfrac{\dfrac{a+b}{J_+(a+b)}\,\mathsf{R}_{J_+}}{h(J_+(a+b))}\,\mathsf{R}_h \quad h(a)}{h(J_+(b))}\,\mathsf{R}_\bullet
$$

An an example of proof rewriting:

$$
\dfrac{\dfrac{\dfrac{a+b \quad c}{a+b+c}\,\mathsf{R}_+}{h(a+b+c)}\,\mathsf{R}_h \quad h(J_+(a))}{h(b+c)}\,\mathsf{R}_\bullet
\quad \Longrightarrow \quad
\dfrac{\dfrac{a+b}{h(a+b)}\,\mathsf{R}_h \quad h(J_+(a)) \quad \dfrac{c}{h(c)}\,\mathsf{R}_h}{h(b+c)}\,\mathsf{R}_\bullet
$$

To prove this result, we consider normal proofs of t which are minimal in size. Then we prove the result by induction on the number of layers. Then it is a series of case study, mainly relying on Lemmas 3, 4, and technical lemmas carefully investigating the cases in which there is a decomposition. We normalize the proofs according to the rules given in Figure 2. These rules are (strongly) terminating (but not confluent). This is our notion of cut elimination.

Before we switch to the proof of this proposition in the next subsections, let us note that theorem 1 is a consequence of the proposition and the following lemma:

Lemma 6 (one-step deducibility). *Given a finite set of terms T, a term t and a function symbol f, it can be decided in polynomial time whether there are terms $t_1, \ldots, t_n \in T$ such that $f(t_1, \ldots, t_n)\downarrow = t$.*

The proof of this lemma relies on standard techniques: if f is not an associative-commutative symbol, then n is fixed and a simple enumeration gives a polynomial algorithm. Otherwise, in all cases, except when $f = \bullet$, only the Abelian group properties of a single symbol have to be considered and the problem amounts to solve a system of linear equations over $\mathbb{Z}$, as already noticed by several authors. When $f = \bullet$, we have that $t = h(t') \bullet t''$ for some terms t' and t''. Let, for $u \in T$, $f_1(u)$, $f_2(u)$ be such that $u = h(f_1(u)) \bullet f_2(u)$,$f_1(u)$ and $f_2(u)$ being possibly empty. The one-step deducibility problem reduces to a system of two systems of linear equations over $\mathbb{Z}$: $\sum_{u \in T} z_u f_1(u) = t'$ and $\prod_{u \in T} f_2(u)^{z_u} = t''$.

Now, the algorithm works as follows. Given T and t, we compute $F(T) \cup F(t)$ (linear time) and then use a fixed point algorithm for the computation of deducible terms in $F(T) \cup F(t)$. Initially, the set D of deducible terms is set to T. Then until a fixed point is reached, add to D the terms in $F(T) \cup F(t)$ that are deducible in one step from terms in D.

This is a polynomial algorithm as one-step deducibility can be checked in polynomial time according to Lemma 6

4.3 Preliminary lemmas

Lemma 7. *If t is obtained by decomposition using $\mathsf{R}_f \in \mathcal{I}_1$, one of the following holds:*

- $t \in \{e_+, e_\star, e_\bullet\}$
- *The premise is $f(t)$ ($f \in \{J_+, J_\star, J_\bullet\}$)*
- $f \in \{\star, +, \bullet\}$ *and there is a premise u such that $t \in \mathrm{DS}_f(u)$.*

Proof. The rule R_f can be a decomposition only when $f \in \{J_\circ, \circ\}$ and $\circ \in \{\star, +, \bullet\}$. If $f = J_\circ$, we are in the second case of the conclusion of the lemma. Only remains to consider $f \in \{\star, +, \bullet\}$. Let then $t_1, \ldots, t_n$ be the premises of the rule and t be the conclusion. Either $t = e_f$ (then we fall into the first case of the conclusion) or else $(f(\mathsf{inv}_f(t), t_1, \ldots, t_n))\downarrow = e_\circ$ and we can apply Lemma 3: $\mathsf{inv}_f(\mathsf{inv}_f(t)) \in \mathrm{DS}_f(t_i)$ for some i, which is the desired result. □

$$\frac{\dfrac{h(t_1)\quad t_2\quad \dots\quad t_n}{h(t_1\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_1\qquad u_2\quad\dots\quad u_m}{h(t_1\star t_2\dots\star t_n\star u_2\dots\star u_m)\downarrow}\,\mathsf{Exp}_1 \Rightarrow \frac{h(t_1)\quad t_2\quad\dots\quad t_n\quad u_2\quad\dots\quad u_m}{h(t_1\star t_2\dots\star t_n\star u_2\dots\star u_m)\downarrow}\,\mathsf{Exp}_1$$

$$\frac{\dfrac{\mathsf{exp}(t_1,t_2)\quad t_3\quad \dots\quad t_n}{\mathsf{exp}(t_1,t_2\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_2\qquad u_3\quad\dots\quad u_m}{\mathsf{exp}(t_1,t_2\star\dots t_n\star u_3\star\dots\star u_m)\downarrow}\,\mathsf{Exp}_2 \Rightarrow \frac{\mathsf{exp}(t_1,t_2)\quad t_3\quad\dots\quad t_n\quad u_3\quad\dots\quad u_m}{\mathsf{exp}(t_1,t_2\star\dots t_n\star u_3\star\dots\star u_m)\downarrow}\,\mathsf{Exp}_2$$

$$\frac{h(t_1)\quad t_2\quad\dots\quad \dfrac{u_1\quad\dots\quad u_m}{(u_1\star\dots\star u_m)\downarrow}\,\mathsf{R}_\star\quad\dots\quad t_n}{h(t_1\star t_2\dots u_1\star\dots u_m\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_1 \Rightarrow \frac{h(t_1)\quad t_2\quad\dots\quad u_1\quad\dots\quad u_m\quad\dots\quad t_n}{h(t_1\star t_2\dots u_1\star\dots u_m\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_1$$

$$\frac{\mathsf{exp}(t_1,t_2)\quad\dots\quad \dfrac{u_1\quad\dots\quad u_m}{(u_1\star\dots\star u_m)\downarrow}\,\mathsf{R}_\star\quad\dots\quad t_n}{\mathsf{exp}(t_1,t_2\star\dots\star u_1\star\dots\star u_m\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_2 \Rightarrow \frac{\mathsf{exp}(t_1,t_2)\quad\dots\quad u_1\quad\dots\quad u_m\quad\dots\quad t_n}{\mathsf{exp}(t_1,t_2\star\dots\star u_1\star\dots\star u_m\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_2$$

$$\frac{\dfrac{t_1\quad t_2}{\mathsf{exp}(t_1,t_2)}\,\mathsf{Exp}_3\qquad t_3\quad\dots\quad t_n}{\mathsf{exp}(t_1,t_2\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_2 \Rightarrow \frac{t_1\qquad \dfrac{t_2\quad\dots\quad t_n}{(t_2\star\dots\star t_n)\downarrow}\,\mathsf{R}_\star}{\mathsf{exp}(t_1,t_2\star\dots\star t_n)\downarrow}\,\mathsf{Exp}_3$$

$$\frac{t_1\quad\dots\quad \dfrac{u_1\quad\dots\quad u_m}{(u_1\circ\dots\circ u_m)\downarrow}\,\mathsf{R}_\circ\quad\dots\quad t_n}{(t_1\circ\dots\circ u_1\circ\dots\circ u_m\circ\dots\circ t_n)\downarrow}\,\mathsf{R}_\circ \Rightarrow \frac{t_1\quad\dots\quad u_1\quad\dots\quad u_m\quad\dots\quad u_n}{(t_1\circ\dots\circ u_1\circ\dots\circ u_m\circ\dots\circ t_n)\downarrow}\,\mathsf{R}_\circ$$

$$\frac{\dfrac{u_1\quad\dots\quad u_m}{(u_1+\dots+u_m)\downarrow}\,\mathsf{R}_+}{h((u_1+\dots+u_m)\downarrow)}\,\mathsf{R}_h \Rightarrow \frac{\dfrac{u_1}{h(u_1)}\,\mathsf{R}_h\quad\dots\quad\dfrac{u_m}{h(u_m)}\,\mathsf{R}_h}{(h(u_1+\dots+u_m))\downarrow}\,\mathsf{R}_\bullet$$

$$\frac{\dfrac{u_1\quad\dots\quad u_m}{(u_1\circ\dots\circ u_m)\downarrow}\,\mathsf{R}_\circ}{J_\circ(u_1\circ\dots\circ u_m)\downarrow}\,\mathsf{R}_{J_\circ} \Rightarrow \frac{\dfrac{u_1}{J_\circ(u_1)\downarrow}\,\mathsf{R}_{J_\circ}\quad\dots\quad\dfrac{u_m}{J_\circ(u_m)\downarrow}\,\mathsf{R}_{J_\circ}}{J_\circ(u_1\circ\dots\circ u_m)\downarrow}\,\mathsf{R}_\circ$$

$$\frac{\dfrac{u}{h(u)}\,\mathsf{R}_h\qquad v_1\ \cdots\ v_n}{h(u\star v_1\star\dots\star v_n)\downarrow}\,\mathsf{Exp}_1 \Rightarrow \frac{\dfrac{u\ v_1\ \cdots\ v_n}{(u\star v_1\star\dots\star v_n)\downarrow}\,\mathsf{R}_\star}{h(u\star v_1\star\dots\star v_n)\downarrow}\,\mathsf{R}_h$$

Fig. 2. Proof normalization rules

Lemma 8. *If t is obtained by a decomposition rule of $\mathcal{I}_2$, then the premises can be written $h(t_1), t_2, \ldots, t_n$, $t = h(u)$ and there an index i such that $t_i = e_\star$ or $u \in \mathrm{DS}_\star(t_i)$.*

Proof. By definition, only Exp_1 can be a decomposition: the premises are $h(t_1)$, $t_2, \ldots, t_n$ and the conclusion is $h(u)$ with $u = t_1 \star \ldots \star t_n\!\downarrow$. Now, if $\textsc{top}(u) \neq \star$, by definition, the rule $\mathsf{R}_\star$ with premises $t_1, \ldots, t_n$ and conclusion u is a decomposition. By Lemma 7, $u = e_\star$ or there is an index i such that $u \in \mathrm{DS}_\star(t_i)$. $\square$

Lemma 9. *For any set of terms in normal form, $F(F(T)) = F(T)$.*

Proof. $F(T) \subseteq F(F(T))$ by definition. For the converse inclusion, first terms in $\mathrm{Sub}(F(T))$ which are not in $\mathrm{Sub}(T)$ are always in $F(T)$. Now, we investigate each other case:

- If $t = h(u)$ with $u \in \mathrm{Sub}(F(T))$ and $\textsc{top}(u) = +$. Then it follows that $u \in \mathrm{Sub}(T)$ or $u = \mathsf{inv}_+(v)$ with $v \in \mathrm{Sub}(T)$, hence $t \in F(T)$
- If $t = h(\mathsf{inv}_+(u))$ with $u \in \mathrm{Sub}(F(T))$ and $\textsc{top}(u) = +$, it is the same as above.
- If $t = \mathsf{inv}_\circ(u)$ with $u \in \mathrm{Sub}(F(T))$ and $\textsc{top}(u) = \circ \in \{\star, +\}$, then either $u \in \mathrm{Sub}(T)$ or $\mathsf{inv}_\circ(u) \in \mathrm{Sub}(T)$. In both cases $t = \mathsf{inv}_\circ(u) \in F(T)$.
- If $t = h(u)$ with $u \in \mathrm{DS}_\circ(v)$, $\textsc{top}(u) = \circ$ and $v \in \mathrm{Sub}(F(T))$ and $\circ \in \{\star, +\}$ then either v or $\mathsf{inv}_\circ(v)$ is in $\mathrm{Sub}(T)$. In the first case we get $t \in F(T)$. In the latter case $\mathsf{inv}_\circ(u) \in \mathrm{DS}_\circ(\mathsf{inv}_\circ(v))$, hence $t = h(\mathsf{inv}_\circ(\mathsf{inv}_\circ(u))) \in F(T)$.
- If $t = \mathsf{inv}_\circ(u)$, $v \in \mathrm{Sub}(F(T))$, $\textsc{top}(v) = \circ \in \{\star, +, \bullet\}$ and $u \in \mathrm{DS}_\circ(v)$, again either $v \in \mathrm{Sub}(T)$ or $\mathsf{inv}_\circ(v) \in \mathrm{Sub}(T)$. In the first case we conclude directly $t \in F(T)$. In the latter case, $\mathsf{inv}_\circ(u) \in \mathrm{DS}_\circ(\mathsf{inv}_\circ(v))$, hence $t \in \mathrm{Sub}(T)$
- The last case is similar to previous ones. $\square$

4.4 Proof of our locality result

We are now able to prove our locality result.

Proposition 2. *The inference system $\mathcal{I}_1 \cup \mathcal{I}_2$ is F-local.*

Proof. We consider a minimal (in terms of size) normal proof of t from the set of hypotheses H. We prove by induction on the proof size that, if the last rule is a composition, then all terms in the proof belong to $F(H) \cup F(t)$ and, if the last rule is a decomposition, then all terms in the proof belong to $F(H)$. In the base case, the proof consists of an axiom only and the result follows. Otherwise, we distinguish cases depending on the last rule used in the proof.

The last rule is R_h. If $\textsc{top}(t) \neq \bullet$, then we simply have to apply the induction hypothesis: $t = h(u)$ and all terms in the proof of u are in $F(H) \cup F(u)$, hence in $F(H) \cup F(t)$.
Now, if $t = h(u)$, then, by proof normalization, u cannot be obtained by R_+. It follows that it must be obtained by decomposition (or possibly R_{J_+}). In any case, $u \in F(H)$ by induction hypothesis and, since $\textsc{top}(u) = +$, $u \in \mathrm{Sub}(H) \cup J_+(\mathrm{Sub}(H))$. It follows that $t, u \in F(H)$.

If the last rule is a composition $\mathsf{R}_\circ$ with $\circ \in \{+, \star, \bullet\}$.

$$\dfrac{\dfrac{\Pi_1}{u_1}\,\mathsf{R}_{f_1} \quad \ldots \quad \dfrac{\Pi_n}{u_n}\,\mathsf{R}_{f_n}}{t}\,\mathsf{R}_\circ$$

Consider the set S of indices i such that $\textsc{top}(u_i) = \circ$ (we may rule out the cases where $u_i = e_\circ$, which correspond to non-minimal proofs). By lemma 4, for every $i \notin S$, either $u_i \in \mathrm{DS}_\circ(t)$ or there is an index j such that $\mathsf{inv}_\circ(u_i) \in \mathrm{DS}_\circ(u_j)$. In the first case $u_i \in \mathrm{Sub}(t)$. In the second case, we claim that if $\circ \in \{\star, +\}$, then u_j must be obtained by decomposition: $f_j \notin \{\circ, J_\circ\}$ by proof normalization and therefore $\textsc{top}(u_j) = \circ$ implies it is obtained by decomposition (this does not hold when $\circ = \bullet$). In case $\circ = \bullet$, either u_j is obtained by decomposition, or $u_j = h(v_j)$ is obtained by R_h and, by proof normalization and since $\textsc{top}(u_j) = \bullet$, v_j must be obtained by decomposition and $\textsc{top}(v_j) = +$.

By induction hypothesis, $u_j \in F(H)$ or $u_j = h(v_j)$ and $v_j \in F(H)$, $\textsc{top}(v_j) = +$, in which case, again $u_j \in F(H)$. And $\mathsf{inv}_\circ(u_i) \in \mathrm{Sub}(u_j)$, hence $u_i \in F(H)$.

To sum up: for every i, either $\textsc{top}(u_i) \neq \circ$ and $u_i \in \mathrm{Sub}(t) \cup F(H)$ or else $\textsc{top}(u_i) = \circ$ and $u_i \in F(H)$. By the induction hypothesis, for every i, all terms in the proof of u_i belong to $F(H)$ or to $F(u_i)$. Hence, by lemma 9, all terms in the proof of t belong to $F(t) \cup F(H)$.

The last rule is $\mathsf{R}_{J_\circ}$. Let $t = J_\circ(u)\!\downarrow$. By proof normalization, u is not obtained by $\mathsf{R}_\circ$ and by minimality, it cannot be obtained by $\mathsf{R}_{J_\circ}$. Then, if $\textsc{top}(u) \in \{\circ, J_\circ\}$, u must be obtained by decomposition or (R_h and $\circ = \bullet$). In both cases $u \in F(H)$ (either the induction hypothesis or the first case above).

Now, if $\textsc{top}(u) \notin \{\circ, J_\circ\}$, $t = J_\circ(u)$ and $u \in \mathrm{Sub}(t)$. We conclude by applying the induction hypothesis.

The last rule is a decomposition $\mathsf{R}_\circ$, $\circ \in \{\star, \bullet, +\}$. Let $t = t_1 \circ \ldots \circ t_n\!\downarrow$. We discard the cases in which $t_i = e_\circ$ for some i (there is a simpler proof). Then the rule being a decomposition, by lemma 7, $t \in \mathrm{Sub}(t_i)$ for some i.

If $t_1, \ldots, t_n \in F(H)$, then by induction hypothesis, we get a proof in which all terms are in $F(H)$. Otherwise, let us assume that some t_j is not in $F(H)$, hence, by induction hypothesis, t_j is obtained by composition.

By contradiction, assume $\textsc{top}(t_j) = \circ$. Then, because it is obtained by composition and because of proof normalization rules, either t_j is obtained by $\mathsf{R}_{J_\circ}$ or $\circ = \bullet$ and t_j is obtained by R_h. In the first case, $t_j = J_\circ(u_j)\!\downarrow$ and, by proof normalization, u_j is not obtained by $\mathsf{R}_\circ$, while $\textsc{top}(u_j) = \circ$. This implies that u_j is obtained by decomposition, and therefore, by induction hypothesis, $u_j \in F(H)$, which in turn contradicts $t_j \notin F(H)$. In the second case, $t_j = h(u_j)$ and $\textsc{top}(u_j) = +$. By proof normalization, u_j cannot be obtained by R_+. It follows that u_j is obtained by decomposition (or else by R_{J_+}). Again, this will yield a contradiction with $t_j \notin F(H)$. Similarly, we rule out $\textsc{top}(t_j) = J_\circ$.

Now, $\textsc{top}(t_j) \notin \{\circ, J_\circ\}$ and, by lemma 3, either $t_j = e_\circ$ (in which case there is a simpler proof) or $t_j = t$ (in which case there is a simpler proof) or else there is an index $k \neq j$ such that $\mathsf{inv}_\circ(t_j) \in \mathrm{DS}_\circ(t_k)$. Moreover, we cannot have $t_k = \mathsf{inv}_\circ(t_j)$: there would be a simpler proof, simply discarding t_j and t_k from the proof of t. Therefore $\textsc{top}(t_k) = \circ$.

Now, we can reason as for t_j: by proof normalization, t_k must have been obtained by decomposition and therefore, by induction hypothesis, $t_k \in F(H)$. Then $t_j \in F(H)$ (since $\mathsf{inv}_\circ(t_j) \circ u \in F(H)$ for some u). This is again a contradiction.

It follows that all t_i's must be in $F(H)$.

The last rule is a decomposition Exp_1.

$$\frac{h(t_1) \;\; t_2 \;\; \ldots \;\; t_n}{h(t_1 \star \ldots \star t_n)\downarrow}$$

and $t = h(u) = h(t_1 \star \ldots \star t_n)\downarrow$ and moreover, $\textsc{top}(u) \neq \star$.

By proof normalization, $h(t_1)$ can only be obtained by $\mathsf{R}_\bullet$, or a decomposition rule distinct from E_1. In all cases, either $h(t_1) \in F(H)$ or else $\textsc{top}(t_1) = +$. Similarly, none of the t_i's ($i \geq 2$) can be obtained by $\mathsf{R}_\star$. Let $u_i = t_i$ if t_i is not obtained by $\mathsf{R}_{J_\star}$ and $u_i = \mathsf{inv}_\star(t_i)$ otherwise. Assume by contradiction that, for some i, $u_i \notin F(H)$. Then, in particular, by induction hypothesis, u_i is obtained by composition. By proof normalization, it cannot be by a rule $\mathsf{R}_\star, \mathsf{R}_{J_\star}$, hence $\textsc{top}(u_i) \neq \star$ and thus $\textsc{top}(t_i) \neq \star$. But then, by lemma 4, either $t_i \in \mathrm{DS}_\star(u)$ or else $\mathsf{inv}_\star(t_i) \in \mathrm{DS}_\star(t_j)$ for some j. We cannot have $t_i = \mathsf{inv}_\star(t_j)$: there would be a simpler proof. Therefore, if $\mathsf{inv}_\star(t_i) \in \mathrm{DS}_\star(t_j)$, we must have $\textsc{top}(t_j) = \star$, hence the corresponding u_j cannot be obtained by composition. It follows that, for every i, either $u_i \in F(H)$ or else $\mathsf{inv}_\star(t_i) \in \mathrm{DS}_\star(t_j)$ and $u_j \in F(H)$. In the latter case, either $u_j = t_j$ and then $\mathsf{inv}_\star(t_i) \in \mathrm{Sub}(t_j)$ implies $t_i \in F(H)$. or else $u_j = \mathsf{inv}_\star(t_j)$ and $t_i \in \mathrm{DS}_\star(u_j) \subseteq F(H)$.

In all cases, for every i, $t_i \in F(H)$.

By lemma 8, there is an index i such that $u \in \mathrm{DS}_\star(t_i)$. Now, $t \neq t_i$ and $t \neq h(t_i)$ (otherwise there is a simpler proof). Hence $\textsc{top}(t_i) = \star$. It follows that $t_i \in \mathrm{Sub}(H) \cup J_\star(\mathrm{Sub}(H))$. Then $u \in \mathrm{Sub}(H) \cup J_\star(\mathrm{Sub}(H))$ and there exists a v such that $u \star v \in \mathrm{Sub}(H)$: $t = h(u) \in F(H)$.

The last rule is a composition Exp_1. We use the same notations as in the previous case. By proof normalization, for every i, $t_i \in F(H)$ or else $\textsc{top}(t_i) \neq \star$. We apply again lemma 4: if $t_i \notin F(H)$, then either $\mathsf{inv}_\star(t_i) \in \mathrm{DS}_\star(t_j)$ for some $j \neq i$ or else $t_j \in \mathrm{DS}_\star(u)$. In the first case, we will have again $t_i \in F(H)$ since $t_j \neq \mathsf{inv}_\star(t_i)$ by minimality and therefore $t_j \in F(H)$. In the end, for every $i > 1$, either $t_i \in F(H)$ or $t_i \in F(t)$.

For $h(t_1)$, the reasoning is the same as in the previous case: either $\textsc{top}(t_1) \neq \star$ or $h(t_1)$ is obtained by decomposition.

The last rule is Exp_2.

$$\frac{\mathsf{exp}(t_1, t_2) \;\; t_3 \;\; \ldots \;\; t_n}{\mathsf{exp}(t_1, t_2 \star \ldots \star t_n)\downarrow}$$

We let $u = t_2 \star \ldots \star t_n\downarrow$
In this case, $t_2, t_3, \ldots, t_n$ play exactly the same roles as $t_1, \ldots$ in the Exp_1 case. It is actually even simpler: $\mathsf{exp}(t_1, t_2)$ cannot be obtained by a composition (by proof normalization), hence $\mathsf{exp}(t_1, t_2) \in \mathrm{Sub}(H)$. Concerning $t_3, \ldots, t_n$, for each index i, either $t_i \in F(H)$ or else $t_i \in \mathrm{DS}_\star(u)$. So, each of the premises is either in $F(H)$ or in $F(t)$ and it suffices to apply the induction hypothesis (together with lemme 9).

The last rule is Exp_3. In this case, the last rule is a composition rule and the two premises are in $\mathrm{Sub}(t)$. It suffices to apply the induction hypothesis. □

5 Conclusion and future works

This paper is only one of the steps on the way of solving the case study. We only showed that, for a passive intruder, the security problem can be solved in PTIME.

The next step (on which we are currently working) is to consider an active intruder. This means solving deducibility constraints modulo the equational theory of the electronic purse.

In this context, we will use properties of the rewrite system, for instance the finite variant property. That is why the rewrite rules are oriented in an unusual way, which introduces some technical complications.

Acknowledgments

We thank an anonymous referee for his comments and we thank Jean-Pierre Jouannaud for introducing us to this area of research.

References

1. M. Arnaud, V. Cortier, and S. Delaune. Combining algorithms for deciding knowledge in security protocols. Research Report 6118, INRIA, Feb. 2007. 28 pages.
2. Y. Chevalier, R. Kuester, M. Rusinowitch, and M. Turuani. An NP decision procedure for protocol insecurity with xor. In Kolaitis [14].
3. Y. Chevalier, R. Küsters, M. Rusinowitch, and M. Turuani. Deciding the security of protocols with Diffie-Hellman exponentiation and products in exponents. In J. Radhakrishnan and P. Pandya, editors, *Proc. FST/TCS, Mumbai*, volume 2914 of *Lecture Notes in Computer Science*, 2003.
4. Y. Chevalier and M. Rusinowitch. Combining intruder theories. In *Proc. ICALP*, volume 3580 of *Lecture Notes in Computer Science*, pages 639–651, 2005.
5. Y. Chevalier and M. Rusinowitch. Hierarchical combination of intruder theories. In *Proc. Rewriting Techniques and Applications*, volume 4098 of *Lecture Notes in Computer Science*, pages 108–122, Seattle, 2006.
6. H. Comon-Lundh and S. Delaune. The finite variant property: How to get rid of some algebraic properties. In J. Giesl, editor, *Proceedings of the 16th International Conference on Rewriting Techniques and Applications (RTA'05)*, volume 3467 of *Lecture Notes in Computer Science*, pages 294–307, Nara, Japan, Apr. 2005. Springer.

7. H. Comon-Lundh and V. Shmatikov. Intruder deductions, constraint solving and insecurity decision in preence of exclusive or. In Kolaitis [14].
8. E. Contejean and C. Marché. CiME: Completion modulo E. In *Proc. Rewriting Techniques and Applications*, volume 1103 of *Lecture Notes in Computer Science*, pages 416–419, 1996.
9. 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.
10. S. Delaune, P. Lafourcade, D. Lugiez, and R. Treinen. Symbolic protocol analysis in presence of a homomorphism operator and *exclusive or*. In M. Buglesi, B. Preneel, V. Sassone, and I. Wegener, editors, *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP'06) — Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 132–143, Venice, Italy, July 2006. Springer.
11. N. Dershowitz and J.-P. Jouannaud. Rewrite systems. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 243–309. North Holland, 1990.
12. D. Dolev and A. Yao. On the security of public key protocols. In *Proc. IEEE Symp. on Foundations of Computer Science*, pages 350–357, 1981.
13. D. Kapur, P. Narendran, and L. Wang. Analyzing protocols that use modular exponentiation: Semantic unification techniques. In *Proc. Rewriting Techniques and Applications*, volume 2706 of *Lecture Notes in Computer Science*, 2003.
14. P. Kolaitis, editor. *18th Annual IEEE Symposium on Logic in Computer Science*, Ottawa, Canada, June 2003. IEEE Computer Society.
15. J. Millen and V. Shmatikov. Symbolic protocol analysis with products and diffie-hellman exponentiation. Invited submission to Journal of Computer Security (selected papers of CSFW-16), 2004.
16. M. Rusinowitch and M. Turuani. Protocol insecurity with finite number of sessions is NP-complete. In *Proc.14th IEEE Computer Security Foundations Workshop*, Cape Breton, Nova Scotia, June 2001.
17. V. Shmatikov. Decidable analysis of cryptographic protocols with products and modular exponentiation. In D. Schmidt, editor, *Proc. European Symposium on Programming (ESOP'04)*, volume 2986 of *Lecture Notes in Computer Science*, pages 355–369. Springer-Verlag, 2004.
18. Security protocols open repository. http://www.lsv.ens-cachan.fr/spore/.