On the Decidability of (ground) Reachability Problems for Cryptographic Protocols (extended version)

Yannick Chevalier, Mounira Kourjieh

On the Decidability of (ground) Reachability Problems for Cryptographic Protocols (extended version)

Yannick Chevalier[1] and Mounira Kourjieh[2]

[1] INRIA Nancy Grand Est, Loria, France, email:Yannick.Chevalier@loria.fr
[2] IRIT, Université de Toulouse, France, email: Mounira.Kourjieh@irit.fr

Abstract. Analysis of cryptographic protocols in a symbolic model is relative to a deduction system that models the possible actions of an attacker regarding an execution of this protocol. We present in this paper a transformation algorithm for such deduction systems provided the equational theory has the finite variant property. the termination of this transformation entails the decidability of the ground reachability problems. We prove that it is necessary to add one other condition to obtain the decidability of non-ground problems, and provide one new such criterion.

1 Introduction

Cryptographic protocols are programs designed to ensure secure electronic communications between participants using an insecure networks. Unfortunately, the existence of cryptographic primitives such as encryption and digital signature is not sufficient to ensure security and several attacks were found on established protocols [16,1]. The most relevant example is the bug of the Needham-Schroeder protocol found by Lowe [24] using a model-checking tool. It took 17 years since the protocol was published to find the attack, a so-called *man-in-the-middle* attack. This situation leads to the development of tools and decision procedures for the formal verification of security protocols. There are several approaches to modelling cryptographic protocols and analysing their security properties: reachability analysis (e.g.: NRL [28]), model checking (FDR [25,26], Mur_φ [30]), modal logic and deduction [10], process calculi like *the spi-calculus* [2], so-called cryptographic proofs ([3]) and others. Here, we use yet another technique, based on the resolution of reachability problems.

Early works on verification of cryptographic protocols studied the standard Dolev-Yao intruder model [32] and the perfect cryptography [21] which states that it is impossible to obtain any information about an encrypted message without knowing the exact key necessary to decrypt this message. Unfortunately, this perfect cryptography assumption has proven too idealistic: there are protocols which can be proven secure under perfect cryptography assumption, but which are in reality insecure since an attacker can use properties of the cryptographic primitives in combinaison with the protocol rules in order to attack protocol.

These properties (so-called algebraic properties) are typically expressed as equational theories. An overview on algebraic properties of well-known cryptographic primitives can be found in [19]. In this paper, we study the class of equational theories represented by a finite convergent rewrite system and having the finite variant property modulo the empty theory [18].

Another point of interest is that an intruder is modelled by a deduction system representing the possible inferences it can make on the messages it knows. A *ground reachability problem* for a given deduction system consists in giving a proof using the permitted deductions of a fact represented by a ground term t from a set of known facts represented by a finite set of terms E. *General reachability problems* are generalisation of the problem in which the goal t has non variables, and the goal is to find a ground substitution σ of these variables such that the instance $t\sigma$ is provable from a finite set of ground terms E. This generalisation consists in providing intermediate steps to solve.

Proof strategy. In [17], H. Comon-Lundh proposes a two-steps strategy to solve general reachability problems, *i.e.* first to solve the ground reachability problems by invoking some locality argument, and then to reduce general reachability problems to ground ones. The method described in this paper roughly follows this line. We employ the finite variant property to reduce reachability problems modulo an equational theory to reachability problems modulo the empty theory. We then partially compute a transitive closure of the possible deductions. We prove that the termination of this computation implies the decidability of the ground reachability problems. We conjecture that the overall construction amounts to proving that the deduction system is F-local [9]. We then give a new criterion that permits us to reduce general reachability problems to ground reachability problems. This criterion is based on counting the number of variables in a reachability problem before and after a deduction is guessed, and is a generalisation of the one employed for the specific case of the Dolev-Yao intruder model. The intuition behind this criterion is that a deduction rule has to provide more relations between existing fact than it introduces new unknown. We give an example showing that such an additional criterion is needed, in the sense that there exists deduction systems on which the saturation algorithm terminates, but for which the general reachability problems are undecidable. Another contribution of this paper is a decidability result to the ground reachability problems for the theory of blind signature [23] using the initial definition of subterm introduced in [5,8], a similar result was given in [4] using an extended definition of subterm. In addition we give a decidability result to the general reachability problems for a class of subterm convergent equational theories, while a more general result was given in [8], the proof given in this paper for our special case is much shorter.

Related works. Several decidability results have been obtained for cryptographic protocols in a similar setting [6,29,7]. These results have been extended to handle algebraic properties of cryptographic primitives [12,13,11,4]. In [5], a decidability result was given to the ground reachability problems in the case of subterm

convergent equational theories. This result was extended in [4] and a decidability result to the ground reachability problems in the case of locally stable AC-convergent equational theories was given. Moreover, again in [4], a decidability result was given to the ground reachability problems for the theory of blind signature [23] while this theory was not included in [5]. To obtain a decidability result for the theory of blind signature, Abadi and Cortier [4] use a new extended definition of subterm. The result obtained in [5] was extended in [8] in different way than in [4] and a decidability result was obtained to the general reachability problem for the class of subterm convergent equational theory. The first result of our paper is a decidability result to the ground reachability problems for a class of equational theories which includes the class studied in [5]. We note that the class studied in [4] is incomparable with ours and we note also that the proof used in [4] to decide the ground reachability problems for the theory of blind signature is different from the ours. Another result of this paper is a decidability result to the general reachability problem for a class of equational theories under some conditions on the deduction systems and the class studied in [8] is incomparable with ours. In [9], a decidability result was given to the general reachability problems under some syntactic conditions on the intruder deduction rules, this result is incomparable with ours.

2 Preliminaries

We now introduce some notations and basic definitions for terms, equational theories and term rewriting systems (the reader may refer to [20] for more details), and then proceed with the definition of the so-called intruder constraints.

2.1 Terms

We assume given a signature $\mathcal{G}$, an infinite set of variables $\mathcal{X}$ and an infinite set of free constants C. The set of terms built with $\mathcal{G}$ and $\mathcal{X}$ is denoted $\mathrm{T}(\mathcal{G},\mathcal{X})$ and its subset of ground terms (terms without variables) $\mathrm{T}(\mathcal{G})$. We denote $\mathrm{Var}(t)$ the set of variables occurring in a term $t \in \mathrm{T}(\mathcal{G},\mathcal{X})$, $|\mathrm{Var}(t)|$ the number of elements in the set $|\mathrm{Var}(t)|$ that is the number of distinct variables occurring in t, $\mathrm{Sub}(t)$ the set of subterms of t and $\mathrm{SSub}(t)$ the set of strict subterms of t. These notations are extended as expected to sets of terms. We denote $t[s]$ a term t that admits s as subterm and $t[s \leftarrow s']$ the term t in which s is replaced by s'.

A substitution σ is an involutive mapping from $\mathcal{X}$ to $\mathrm{T}(\mathcal{G},\mathcal{X})$ such that $\mathrm{Supp}(\sigma) = \{x \,|\, \sigma(x) \neq x\}$, the *support* of σ, is a finite set. The application of a substitution σ to a term t (resp. a set of terms E) is denoted $t\sigma$ (resp. $E\sigma$). A substitution σ is *ground* w.r.t. $\mathcal{G}$ if the image of $\mathrm{Supp}(\sigma)$ is included in $\mathrm{T}(\mathcal{G})$.

We recall in the following the definition of *reduction order*:

Definition 1. *Let $\mathcal{G}$ be a signature and $\mathcal{X}$ be an infinite set of variables. A strict order $\succ$ on $\mathrm{T}(\mathcal{G},\mathcal{X})$ is called a rewrite order iff it is*

1. ***compatible*** *with $\mathcal{G}$-function symbols: for all $s, s' \in \mathrm{T}(\mathcal{G}, \mathcal{X})$ and all $f \in \mathcal{G}$ with arity $n \geq 0$, $t_1 \succ t_2$ implies*

$$f(t_1, \ldots, t_{i-1}, s, t_{i+1}, \ldots, t_n) \succ f(t_1, \ldots, t_{i-1}, s', t_{i+1}, \ldots, t_n)$$

   for all i, $1 \leq i \leq n$, and all $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n \in \mathrm{T}(\mathcal{G}, \mathcal{X})$.
2. ***closed under substitutions:*** *for all $s, s' \in \mathrm{T}(\mathcal{G}, \mathcal{X})$ and all substitutions σ, $s \succ s'$ implies $\sigma(s) \succ \sigma(s')$.*

*A **reduction order** is a well-founded rewrite order.*

We consider a reduction order $\succ$ over $\mathrm{T}(\mathcal{G}, \mathcal{X})$ total over ground terms. We denote $\succeq$ the relation between terms such that $t_1 \succeq t_2$ iff $t_1 \succ t_2$ or $t_1 = t_2$ for $t_1, t_2 \in \mathrm{T}(\mathcal{G}, \mathcal{X})$.

A rewriting system $\mathcal{R}$ is a finite set of couples $(l, r) \in \mathrm{T}(\mathcal{G}, \mathcal{X})^2$, where each couple is called a *rewriting rule* and is denoted $l \to r$. The rewriting relation $\to_{\mathcal{R}}$ between terms is defined by $t \to_{\mathcal{R}} t'$ if there exists $l \to r \in \mathcal{R}$ and a substitution σ such that $l\sigma = s$ and $r\sigma = s'$, $t = t[s]$ and $t' = t[s \leftarrow s']$. A rewriting system is *terminating* if for all terms t there is no infinite sequence of rewriting starting from t. It is *convergent* if it has moreover the *confluence* property: every sequence of rewriting ends in the same term denoted $(t)\downarrow_{\mathcal{R}}$, or simply $(t)\downarrow$ if $\mathcal{R}$ is clear from the context. We say that a term t is in *normal form* if $t = (t)\downarrow_{\mathcal{R}}$. A substitution σ is in normal form if for all $x \in \mathrm{Supp}(\sigma)$, the term $\sigma(x)$ is in normal form. Given a substitution σ, we denote $(\sigma)\downarrow_{\mathcal{R}}$ the substitution such that, for all $x \in \mathrm{Supp}(\sigma)$ we have $(x\sigma)\downarrow_{\mathcal{R}} = x(\sigma)\downarrow_{\mathcal{R}}$.

An equational theory $\mathcal{H}$ is a congruence relation on terms in $\mathrm{T}(\mathcal{G}, \mathcal{X})$. We denote $t =_{\mathcal{H}} t'$ the fact that the term t and t' are identified by $\mathcal{H}$. We say that $\mathcal{H}$ is *generated* by a convergent rewriting system $\mathcal{R}$ if $t =_{\mathcal{H}} t'$ iff $(t)\downarrow_{\mathcal{R}} = (t')\downarrow_{\mathcal{R}}$.

2.2 Unification systems

Definition 2. *(Unification systems) Let $\mathcal{H}$ be an equational theory. A $\mathcal{H}$-*unification system *$\mathcal{S}$ is a finite set of pairs of terms in $\mathrm{T}(\mathcal{G}, \mathcal{X})$ denoted by $\left\{u_i \stackrel{?}{=}_{\mathcal{H}} v_i\right\}_{i \in \{1,\ldots,n\}}$. It is satisfied by a substitution σ, and we note $\sigma \models_{\mathcal{H}} \mathcal{S}$, if for all $i \in \{1, \ldots, n\}$ we have $u_i\sigma =_{\mathcal{H}} v_i\sigma$. In this case we call σ a* solution *or a* unifier *of $\mathcal{S}$.*

When $\mathcal{H}$ is generated by a convergent rewriting system $\mathcal{R}$, considering a bottom-up normalisation shows that if σ is a solution of a $\mathcal{H}$-unification system, then $(\sigma)\downarrow$ is also a solution of the same unification system. A top-down normalisation on solutions also demonstrates that we can assume that terms in a unification system are in normal form. Accordingly we will consider in this paper only solutions in normal form of unification systems in normal form. A unifier σ is more general than a unifier τ if there exists a substitution θ such that $\sigma\theta = \tau$. A *complete set of unifiers* of a $\mathcal{H}$-unification system $\mathcal{S}$ is a set Σ of unifiers of $\mathcal{S}$ such that, for any unifier τ of $\mathcal{S}$, there exists $\sigma \in \Sigma$ which is more

general than τ. The unifier τ is a *most general unifier* of $\mathcal{S}$ if the substitution θ in the preceding equation is a variable renaming. We denote $mgu(\mathcal{S})$ the set of most general unifiers modulo $\mathcal{H}$ of a unification system $\mathcal{S}$. In the context of unification modulo an equational theory, standard (or syntactic) unification will also be called unification in the empty theory. In this case, it is well-known that there exists a unique most general unifier of a set of equations. This unifier is denoted $mgu(\mathcal{S})$, or $mgu(s,t)$ in the case $\mathcal{S} = \left\{ s \stackrel{?}{=}_\emptyset t \right\}$.

Finite Variant Property. We will abusively write that an equational theory $\mathcal{H}$ has the *finite variant property* if the couple $(\mathcal{H}, \emptyset)$ has the finite variant property in the notation of [18]. Let us now formally state the definition of this property in this case, simplified using the Lemma 3 and the Theorem 1 of [18].

Definition 3. *(Finite Variant Property) A theory $\mathcal{H}$ has the* finite variant property *if, for any term t, one can compute a finite set of substitutions $\theta_1, \ldots, \theta_n$ (the variant substitutions) such that, for any substitution σ in normal form there exists $i \in \{1, \ldots, n\}$ and a substitution σ' in normal form such that $\sigma = \theta_i\sigma'$ and $(t\sigma)\!\downarrow = (t\theta_i)\!\downarrow\sigma'$. The terms $(t\theta_i)\!\downarrow$ are called the* variants *of t.*

Examples of equational theories having the finite variant property are those defined by a convergent rewriting system and such that either basic narrowing [22] terminates or the rewriting system is optimally reducing [31].

The finite variant property ensures that it is possible to compute a complete set of most general unifiers between two terms t and t'. Indeed, it suffices to compute for these two terms the respective sets of variant substitutions $\{\theta_i\}_{i\in\{1,\ldots,m\}}, \{\theta'_j\}_{j\in\{1,\ldots,n\}}$, and to (try to) unify in the empty theory every pair of terms $(t\theta_i)\!\downarrow \stackrel{?}{=} (t'\theta'_j)\!\downarrow$.

In the rest of this paper we will consider equational theories $\mathcal{H}$ having the finite variant property and generated by a convergent rewriting system $\mathcal{R}$.

2.3 Deduction systems

The notions that we give here have been defined in [15]. These definitions have since been generalised to consider a wider class of intruder deduction and constraint systems [14]. Although this general class encompasses all deduction and constraint systems given in this paper, we have preferred to give the simpler definitions from [15] which are sufficient for stating our problem. We will refer, without further justifications, to the model of [14] as *extended* deduction systems. The constraint systems considered and defined here correspond to symbolic derivations [14] in which a most general unifier of the unification system has been applied on the output messages (for Def. 6) and on input variables (for the extended constraint systems).

In the context of a security protocol (see *e.g.* [28] for a brief overview), we model messages as ground terms and intruder deduction rules as rewriting rules

on sets of messages representing the knowledge of an intruder. The intruder derives new messages from a given (finite) set of messages by applying deduction rules. Since we assume some equational axioms $\mathcal{H}$ are satisfied by the function symbols in the signature, all these derivations have to be considered *modulo* the equational theory $\mathcal{H}$ generated by $\mathcal{R}$.

Definition 4. *A* deduction system $\mathcal{I}$ *is given by a triple* $\langle \mathcal{G}, \mathcal{L}, \mathcal{H} \rangle$ *where* $\mathcal{G}$ *is a signature,* $\mathcal{L}$ *is a set of deduction rules* $l \twoheadrightarrow r$*, where* l *a set of terms in* $\mathrm{T}(\mathcal{G}, \mathcal{X})$ *and* r *a term in* $\mathrm{T}(\mathcal{G}, \mathcal{X})$*, and* $\mathcal{H}$ *is an equational theory.*

Each rule $l \twoheadrightarrow r$ in $\mathcal{L}$ defines a deduction relation $\twoheadrightarrow_{l \twoheadrightarrow r}$ between finite sets of terms. Given two finite sets of terms E and F we have $E \twoheadrightarrow_{l \twoheadrightarrow r} F$ if and only if there exits a substitution σ, such that $l\sigma =_{\mathcal{H}} l'$, $r\sigma =_{\mathcal{H}} r'$, $l' \subseteq E$ and $F = E \cup \{r'\}$. We denote $\twoheadrightarrow_{\mathcal{I}}$ the union of the relations $\twoheadrightarrow_{l \twoheadrightarrow r}$ for all $l \twoheadrightarrow r$ in $\mathcal{L}$ and by $\twoheadrightarrow^*_{\mathcal{I}}$ the transitive closure of $\twoheadrightarrow_{\mathcal{I}}$. Note that, given sets of terms E, E', F and F' such that $E =_{\mathcal{H}} E'$ and $F =_{\mathcal{H}} F'$ by definition we have $E \twoheadrightarrow_{\mathcal{I}} F$ iff $E' \twoheadrightarrow_{\mathcal{I}} F'$. We simply denote by $\twoheadrightarrow$ the relation $\twoheadrightarrow_{\mathcal{I}}$ when there is no ambiguity about $\mathcal{I}$.

We recall that $\succeq$ is the extension of the reduction order $\succ$ defined over $\mathrm{T}(\mathcal{G}, \mathcal{X})$.

Definition 5. *A deduction rule* $l \twoheadrightarrow r$ *is a* decreasing *rule if there is a term* $s \in l$ *such that* $s \succeq r$ *and it is* increasing *otherwise.*

From now, if $\mathcal{L}$ is the set of deduction rules, we denote by $\mathcal{L}_{inc}$ the set of increasing rules and by $\mathcal{L}_{dec}$ the set of decreasing rules. By definition of *increasing* and *decreasing* rules, we have $\mathcal{L} = \mathcal{L}_{inc} \cup \mathcal{L}_{dec}$.

A *derivation* D of length n, $n \geq 0$, is a sequence of steps of the form $E_0 \twoheadrightarrow_{\mathcal{I}} E_0, t_1 \twoheadrightarrow_{\mathcal{I}} \cdots \twoheadrightarrow_{\mathcal{I}} E_n$ with finite sets of terms $E_0, \ldots E_n$, and terms $t_1, \ldots, t_n$, such that $E_i = E_{i-1} \cup \{t_i\}$ for every $i \in \{1, \ldots, n\}$. The term t_n is called the *goal* of the derivation. We let $\mathrm{trace}(D)$ be the set of terms constructed during the derivation D, $\mathrm{trace}(D) = E_0 \cup \{t_1, \ldots, t_n\}$. We define $\overline{E}^{\mathcal{I}}$ to be equal to the set of terms that can be deduced from E, $\overline{E}^{\mathcal{I}} = \{t \text{ s.t. } E \ldots^*_{\mathcal{I}} E' \text{ and } t \in E'\}$. If there is no ambiguity on the deduction system $\mathcal{I}$ we write $\overline{E}$ instead of $\overline{E}^{\mathcal{I}}$.

2.4 Constraint systems

We now introduce the constraint systems to be solved for checking protocols. It is presented in [15] how these constraint systems permit to express the reachability of a state in a protocol execution.

Definition 6. *($\mathcal{I}$-Constraint systems) Let* $\mathcal{I} = \langle \mathcal{G}, \mathcal{L}, \mathcal{H} \rangle$ *be a deduction system. An* $\mathcal{I}$-constraint system $\mathcal{C}$ *is denoted* $((E_i \rhd v_i)_{i \in \{1,\ldots,n\}}, \mathcal{S})$ *and is defined by a sequence of pairs* $(E_i, v_i)_{i \in \{1,\ldots,n\}}$ *with* $v_i \in \mathcal{X}$*,* $E_i \subseteq \mathrm{T}(\mathcal{G}, \mathcal{X})$*,* $E_i \subseteq E_{i+1}$ *and* $\mathrm{Var}(E_i) \subseteq \{v_1, \ldots, v_{i-1}\}$ *for* $i \in \{1, \ldots, n\}$*, and by an* $\mathcal{H}$*-unification system* $\mathcal{S}$*.*

An $\mathcal{I}$-Constraint system $\mathcal{C}$ *is satisfied by a substitution* σ *if for all* $i \in \{1, \ldots, n\}$ *we have* $v_i\sigma \in \overline{E_i\sigma}^{\mathcal{I}}$ *and if* $\sigma \models_{\mathcal{H}} \mathcal{S}$*. We denote that a substitution* σ *satisfies a constraint system* $\mathcal{C}$ *by* $\sigma \models_{\mathcal{I}} \mathcal{C}$*.*

Constraint systems are denoted by $\mathcal{C}$ and decorations thereof. Note that if a substitution σ is a solution of a constraint system $\mathcal{C}$, by definition of deduction rules and unification systems the substitution $(\sigma)\downarrow$ is also a solution of $\mathcal{C}$. In the context of cryptographic protocols the inclusion $E_{i-1} \subseteq E_i$ means that the knowledge of an intruder does not decrease as the protocol progresses: after receiving a message a honest agent will respond to it, this response can then be added to the knowledge of the intruder who listens to all communications. The condition on variables stems from the fact that a message sent at step i must be built from previously received messages recorded in the variables $v_j, j < i$, and from the initial knowledge (set of ground terms) of the honest agents. Our goal is to solve the following decision problem.

$\mathcal{I}$-Reachability Problem

Input: An $\mathcal{I}$-constraint system $\mathcal{C}$.

Output: SAT iff there exists a substitution σ such that $\sigma \models_{\mathcal{I}} \mathcal{C}$.

3 Saturation

In the rest of this paper, we suppose that $\mathcal{I}_0 = \langle \mathcal{G}, \mathcal{L}_0, \mathcal{H} \rangle$ is an initial deduction system. *We assume that $\mathcal{L}_0$ is the union of rules $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n)$ for some function symbols $f \in \mathcal{G}$.*

Let $\mathcal{H}$ be an equational theory having the finite variant property and generated by a convergent rewriting system $\mathcal{R}$. The saturation of the set of deduction rules $\mathcal{L}_0$ defined modulo the equational theory $\mathcal{H}$ is the output of the application of the saturation algorithm given by the following two steps:

- *Step 1:* Anticipating the application of rules of $\mathcal{L}_0$ on ground terms in normal form, we define the set $\mathcal{L}$ of rules "in normal form":

$$\mathcal{L} = \bigcup_{\substack{x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n) \in \mathcal{L}_0 \\ \theta \text{ variant subsitution of } f(x_1, \ldots, x_n)}} x_1\theta, \ldots, x_n\theta \twoheadrightarrow (f(x_1, \ldots, x_n)\theta)\downarrow$$

  This union is over finite sets thanks to the finiteness of $\mathcal{L}_0$ and to the finite variant property.
- *Step 2:* Start with $\mathcal{L}' = \mathcal{L}$, repeat the rule given in Figure 1 until no new rule can be added.

$$\frac{l_1 \twoheadrightarrow r_1 \in \mathcal{L}'_{inc} \; ; \quad l_2, s \twoheadrightarrow r_2 \in \mathcal{L}'}{\mathcal{L}' \leftarrow \mathcal{L}' \cup \{(l_1, l_2 \twoheadrightarrow r_2)\sigma\}} \; \begin{array}{c} s \notin \mathcal{X} \\ \sigma = mgu_{\emptyset}(r_1, s) \end{array}$$

Fig. 1. closure rule.

We define two new deduction systems, corresponding each to one step of the saturation algorithm, $\mathcal{I} = \langle \mathcal{G}, \mathcal{L}, \emptyset \rangle$ and $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$. Since in the first step we consider all possible variants of all possible deduction rules, we have:

Lemma 1. *Let E and F be two sets of ground terms in normal form we have: $E \twoheadrightarrow_{\mathcal{I}_0} F$ iff $E \twoheadrightarrow_{\mathcal{I}} F$.*

Proof. Let E and F be two sets of ground terms in normal form and assume there is a rule $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n) \in \mathcal{L}_0$ such that $E \twoheadrightarrow_{x_1,\ldots,x_n \twoheadrightarrow f(x_1,\ldots,x_n)} F$. By definition there exists a ground substitution σ in normal form such that $(x_1, \ldots, x_n)\sigma \subseteq E$ and $F = E \cup \{(f(x_1, \ldots, x_n)\sigma)\!\downarrow\}$. Due to the finite variant property, there exists a variant substitution θ of $f(x_1, \ldots, x_n)$ and a ground normal substitution σ' such that $(f(x_1, \ldots, x_n)\sigma)\!\downarrow = (f(x_1, \ldots, x_n)\theta)\!\downarrow\sigma'$ and $\sigma = \theta\sigma'$. The rule $Img(\theta) \twoheadrightarrow (f(x_1, \ldots, x_n)\theta)\!\downarrow$ was added to $\mathcal{L}$ by Step 1 this implies that $E \twoheadrightarrow_{\mathcal{I}} F$. To prove the converse, notice that if $(x_1, \ldots, x_n)\theta \twoheadrightarrow (f(x_1, \ldots, x_n)\theta)\!\downarrow$ can be applied with the normal ground substitution σ' on E, then the rule $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n)$ can be applied with the ground substitution $\sigma = (\theta\sigma')\!\downarrow$ on E. □

Also, the computation of Step. 2 is correct and complete in the following sense.

Lemma 2. *For any set of ground terms E in normal form and any ground term t in normal form we have: $t \in \overline{E}^{\mathcal{I}}$ if and only if $t \in \overline{E}^{\mathcal{I}'}$.*

Proof. The direct implication is trivial since $\mathcal{L}'$ is initialised with $\mathcal{L}$. Let us prove the converse implication. Assume that there exists a $\mathcal{I}'$-derivation starting from E of goal t. Let us define an arbitrary total order on the rules of $\mathcal{L}$, and we extend this order to rules of $\mathcal{L}' \setminus \mathcal{L}$ as follows: rules of $\mathcal{L}$ are smaller than the rules of $\mathcal{L}' \setminus \mathcal{L}$ and rules of $\mathcal{L}' \setminus \mathcal{L}$ are ordered according to the order of their construction during the saturation. Let $\mathrm{M}(D)$ be the multiset of rules applied in D. Let $\Omega(E, t) = \left\{ D \mid D : E \twoheadrightarrow^*_{\mathcal{I}'} F \ni t \right\}$. By construction, the ordering on rules is total and well-founded, and thus the pre-ordering on derivations in $\Omega(E, t)$ is also total and well-founded. Since $t \in \overline{E}^{\mathcal{I}'}$, we have $\Omega(E, t) \neq \emptyset$, and thus $\mathrm{M}(\Omega(E, t))$ has a minimum element which is reached. Let D be a derivation in $\Omega(E, t)$ having the minimum $\mathrm{M}(D)$, and let us prove that D employs only rules in $\mathcal{L}$. By contradiction, assume that D uses a rule $l \twoheadrightarrow r \in \mathcal{L}' \setminus \mathcal{L}$ applied with a ground substitution σ on a set F. Since $l \twoheadrightarrow r \notin \mathcal{L}$, it has been constructed by closure rule. Thus, there exists two rules $l_1 \twoheadrightarrow r_1 \in \mathcal{L}'_{inc}$ and $l_2 \twoheadrightarrow r_2 \in \mathcal{L}'$, a term $s \in l_2 \setminus \mathcal{X}$ such that s and r_1 are unifiable, $\alpha = mgu(s, r_1)$, $l = (l_1, l_2 \setminus s)\alpha$ and $r = r_2\alpha$. Replacing the application of the rule $l \twoheadrightarrow r$ by two steps applying first the rule $l_1 \twoheadrightarrow r_1$ and then $l_2 \twoheadrightarrow r_2$ yields another derivation D'. Since $l \twoheadrightarrow r$ must have an order bigger than the order of $l_1 \twoheadrightarrow r_1$ and $l_2 \twoheadrightarrow r_2$ and the last two rules are in $\mathcal{L}'$, we deduce that $D' \in \Omega(E, t)$ and $\mathrm{M}(D') < \mathrm{M}(D)$ which contradicts the minimality of $\mathrm{M}(D)$. □

Let E (resp. t) be a set of terms (resp. a term) in normal form and let D be a derivation starting from E of goal t, $D : E = E_0 \twoheadrightarrow E_0, t_1 \twoheadrightarrow \ldots \twoheadrightarrow E_{n-2}, t_{n-1} \twoheadrightarrow E_{n-1}, t$. The derivation D is *well-formed* if for all rules $l \twoheadrightarrow r$ applied with substitution σ, for all $u \in l \setminus \mathcal{X}$ we have either $u\sigma \in E$ or $u\sigma$ was deduced by a former decreasing rule. The following lemma is a consequence of the computation of the closure. Notice that we do not assume here, nor afterward unless stated, that the saturation terminates.

Lemma 3. *Let E (resp. t) be a set of terms (resp. a term) in normal form such that $t \in \overline{E}^{\mathcal{I}'}$. For all $\mathcal{I}'$-derivations D starting from E of goal t we have either D is well-formed or there is another $\mathcal{I}'$-derivation D' starting from E of goal t such that* $\mathrm{trace}(D) = \mathrm{trace}(D')$ *and D' is well-formed.*

Proof. We have $t \in \overline{E}^{\mathcal{I}'}$ implies that the set $\Omega(E, t)$ of $\mathcal{I}'$-derivations starting from E of goal t is not empty. Let $D \in \Omega(E, t)$, $D : E = E_0 \twoheadrightarrow E_1 \twoheadrightarrow \ldots \twoheadrightarrow E_{n-1}, t$, we denote $l_i \twoheadrightarrow r_i$ the rule applied at step i with the substitution σ_i and suppose that D is not well-formed. Let us (pre-)order derivations in $\Omega(E, t)$ with a measure M such that $\mathrm{M}(D')$ for a derivation D' is a multiset of integers constructed as follows: starting with $\mathrm{M}(D') = \emptyset$, for all steps k, $1 \leq k \leq n$, for all terms $u \in l_k\sigma_k$ obtained by former increasing rule, add k to $\mathrm{M}(D')$. Since this pre-order is well-founded, there exists a derivation $d \in \Omega(E, t)$ such that $\mathrm{M}(d)$ is minimum and $\mathrm{trace}(d) = \mathrm{trace}(D)$. Let us prove that d is well-formed. By contradiction, assume that d is not well-formed and let j be the first step in d such that $l_j \twoheadrightarrow r_j$ is the rule applied with substitution σ_j and there is a term $u \in l_j \setminus \mathcal{X}$ obtained by a former increasing rule, let $l_h \twoheadrightarrow r_h$ be this rule. Since $l_h \twoheadrightarrow r_h \in \mathcal{L}'_{inc}$ and $u \notin \mathcal{X}$, *Closure* can be applied on $l_h \twoheadrightarrow r_h$ and $l_j \twoheadrightarrow r_j$ and the resulting rule can be applied at step j instead of $l_j \twoheadrightarrow r_j$ yielding also E_j. Let d' be the derivation obtained after this replacement, $d' \in \Omega(E, t)$ and $\mathrm{trace}(d') = \mathrm{trace}(d)$. Since $h < j$ and by definition of M, we have $\mathrm{M}(d') < \mathrm{M}(d)$ which contradicts the minimality of $\mathrm{M}(d)$. We deduce that d is well-formed and then we have the lemma.

□

4 Reachability problems

4.1 Presentation of the algorithm and pre-computation

This section is devoted to the presentation of an algorithm for solving Reachability Problems and to a proof scheme of its completeness, correctness and termination. In this section, we denote by $\mathcal{I}_0 = \langle \mathcal{G}, \mathcal{L}_0, \mathcal{H} \rangle$ the initial deduction system and by $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$ the saturated deduction system. From now, we suppose that $\mathcal{L}'$ is finite and we recall that $\mathcal{L}'$ is partitioned into two disjoint sets of deduction rules $\mathcal{L}'_{inc}$ and $\mathcal{L}'_{dec}$ (by definition of *increasing* and *decreasing* rules). The algorithm comprises two steps, and is depicted in Fig. 2

Resolution($\mathcal{C}^0$)

We let $\mathcal{C}^0 = ((E_i^0 \rhd v_i^0)_{i\in\{1,\ldots,n\}}, \mathcal{S}^0)$ be an $\mathcal{I}_0$-constraint system.

Step 1. Guess a finite variant substitution θ for all terms of $\mathcal{C}^0$, apply θ on these terms and normalise them then solve the obtained unification system. Finally, apply the obtained solution α on the constraints. In the sequel we will abuse notations and denote the obtained constraint system $\mathcal{C} = (E_i \rhd t_i)_{i\in\{1,\ldots,n\}}$, where $t_i = (v_i^0\theta)\!\downarrow\!\alpha$ and $E_i = (E_i^0\theta)\!\downarrow\!\alpha$.

Step 2. Apply non-deterministically the transformation rules of Fig. 3

Step 3. If a solved form is reached, return SAT, else return FAIL.

Fig. 2. Algorithm for solving constraint systems.

Remarks.

Solved form. A constraint system $\mathcal{C}$ as denoted at the end of the first step is in *solved form* if for all constraints $E \rhd t \in \mathcal{C}$ we have $t \in \mathcal{X}$. Every constraint system in solved form has at least one solution [6].

Computation of the finite variants substitutions. Given $\mathcal{C}^0 = ((E_i^0 \rhd v_i^0)_{1\leq i\leq n}, \mathcal{S}^0)$, and let T be a n-uplet containing terms appearing in $\mathcal{C}^0$, $T = \langle u_1, \ldots, u_n\rangle$. Due to the finite variant property, T has finite set of variant substitutions. We choose a variant substitution θ among the possible ones.

Justification of the first step. Let σ be a normal solution of the original constraint system. The first step will non-deterministically transform terms of $\mathcal{C}$, $u_1, \ldots, u_n$, into terms $u_1^0, \ldots, u_n^0$ such that, according to definition 3 we will have $\left\{(u_i\sigma)\!\downarrow = u_i^0\sigma'\right\}_{1\leq i\leq n}$ for a normal substitution σ'. It is easily verified that the first step always terminates.

We prove below that there exists a solution to the original $\mathcal{I}_0$-constraint system $\mathcal{C}^0$ *iff* there exists a solution to one of the possible constraint systems computed in the first step for the $\mathcal{I}'$ deduction system.

Lemma 4. *(Completeness) Let $\mathcal{C}^0$ be an $\mathcal{I}_0$-constraint system. If $\mathcal{C}^0$ is $\mathcal{I}_0$-satisfiable, there exists a constraint system $\mathcal{C}$ in the output of Step 1. such that $\mathcal{C}$ is $\mathcal{I}'$-satisfiable.*

PROOF. We have $\mathcal{C}^0 = ((E_i^0 \rhd v_i^0)_{i\in\{1,\ldots,n\}}, \mathcal{S}^0)$. Let σ be a substitution in normal form such that $\sigma \models_{\mathcal{I}_0} \mathcal{C}^0$. This implies that $(v_i^0\sigma)\!\downarrow \in \overline{(E_i^0\sigma)\!\downarrow}^{\mathcal{I}_0}$ for $i \in \{1, \ldots, n\}$ and thus, by lemmas 1 and 2, $(v_i^0\sigma)\!\downarrow \in \overline{(E_i^0\sigma)\!\downarrow}^{\mathcal{I}'}$ for $i \in \{1, \ldots, n\}$. We have also $(s^0\sigma)\!\downarrow = (s'^0\sigma)\!\downarrow$ for all equations $s^0 \stackrel{?}{=} s'^0 \in \mathcal{S}^0$. By definition 3, there exists a variant substitution θ of the terms in $\mathcal{C}^0$ and a substitution σ' in normal form such that for each term $u \in \mathcal{C}$, we have $(u\sigma)\!\downarrow = (u\theta)\!\downarrow\sigma'$. This implies that $(v_i^0\theta)\!\downarrow\sigma' \in \overline{(E_i^0\theta)\!\downarrow\sigma'}^{\mathcal{I}'}$ for $i \in \{1, \ldots, n\}$ and $(s^0\theta)\!\downarrow\sigma' = (s'^0\theta)\!\downarrow\sigma'$ for all equations $s^0 \stackrel{?}{=} s'^0 \in \mathcal{S}^0$. The unification system $(\mathcal{S}^0\theta)\!\downarrow$ has solution (σ'), let μ

be its most general unifier, we have $\sigma' = \mu\alpha$ for some substitution α and $\alpha \models_{\mathcal{I}'} (E_i^0\theta)\downarrow\mu \rhd (v_i^0\theta)\downarrow\mu$ for $i \in \{1,\ldots,n\}$. The constraint system $\mathcal{C} = (((E_i^0\theta)\downarrow\mu \rhd (v_i^0\theta)\downarrow\mu)_{i\in\{1,\ldots,n\}})$ is a possible output of Step 1 and it is $\mathcal{I}'$-satisfiable. □

Lemma 5. *(Correctness) Let $\mathcal{C}^0$ (resp. $\mathcal{C}$) be a $\mathcal{I}_0$- (resp. $\mathcal{I}'$-) constraint system. Assume that $\mathcal{C}$ is obtained from $\mathcal{C}^0$ by applying Step 1. If $\mathcal{C}$ is satisfiable then so is $\mathcal{C}^0$.*

Proof. Let $\mathcal{C}^0$ (resp. $\mathcal{C}$) be a $\mathcal{I}_0$- (resp. $\mathcal{I}'$-) constraint system and assume that $\mathcal{C}$ is obtained from $\mathcal{C}^0$ by applying Step 1. This implies that $\mathcal{C}^0 = ((E_i^0 \rhd v_i^0)_{i\in\{1,\ldots,n\}}, \mathcal{S}^0)$ and $\mathcal{C} = (((E_i^0\theta)\downarrow\mu \rhd (v_i^0\theta)\downarrow\mu)_{i\in\{1,\ldots,n\}})$ while θ is a variant substitution of the terms of $\mathcal{C}^0$ and μ is the most general unifier of the unification system $(\mathcal{S}^0\theta)\downarrow$ obtained from $\mathcal{S}^0$ by applying the variant substitution θ on the terms of $\mathcal{S}^0$ and then normalising these terms. Since $\mathcal{C}$ is $\mathcal{I}'$-satisfiable there exists a normal substitution σ such that $(v_i^0\theta)\downarrow\mu\sigma \in \overline{(E_i^0\theta)\downarrow\mu\sigma}^{\mathcal{I}'}$ and thus $(v_i^0\theta\mu\sigma)\downarrow \in \overline{(E_i^0\theta\mu\sigma)\downarrow}^{\mathcal{I}_0}$ (lemmas 1 and 2). We conclude that $(\theta\mu\sigma)\downarrow \models_{\mathcal{I}_0} \mathcal{C}^0$. □

4.2 Transformation in solved form

Let $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$ be the deduction system resulting from the application of the saturation algorithm. In the rest of this paper, we denote by $l_x, l_1, \ldots, l_n \twoheadrightarrow r$ a $\mathcal{L}'$-rule such that l_x is a finite set of variables and $\{l_1, \ldots, l_n\}$ is a finite set of non-variable terms. Unless otherwise specified, $\mathcal{I}'$ is the deduction system implicit in all notations.

In the rest of this section, we prove a *progress* property: If a satisfiable constraint system is not in solved form, then a rule of Fig. 3 can be applied on it to yield another satisfiable constraint system. We will give conditions in the next section ensuring the termination of the application of these rules.

Unif :
$$\frac{\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta}{(\mathcal{C}_\alpha, \mathcal{C}_\beta)\sigma} \quad \begin{array}{c} u \in E \setminus \mathcal{X},\ t \notin \mathcal{X}, \\ \sigma = mgu(u,t) \end{array}$$

Reduce 1 :
$$\frac{\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta}{(\mathcal{C}_\alpha, (E \rhd y)_{y\in l_x}, \mathcal{C}_\beta)\sigma} \quad \begin{array}{c} l_x, l_1, \ldots, l_n \twoheadrightarrow r \in \mathcal{L}'_{inc} \text{ and } t \notin \mathcal{X} \\ e_1, \ldots, e_n \in E \setminus \mathcal{X} \text{ and } \sigma = mgu(\left\{e_i \stackrel{?}{=} l_i\right\}_{1\le i\le n}, r \stackrel{?}{=} t) \end{array}$$

Reduce 2 :
$$\frac{\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta}{(\mathcal{C}_\alpha, (E \rhd y)_{y\in l_x}, E \cup r \rhd t, \mathcal{C}'_\beta)\sigma} \quad \begin{array}{c} l_x, l_1, \ldots, l_n \twoheadrightarrow r \in \mathcal{L}'_{dec} \text{ and } t \notin \mathcal{X} \\ e_1, \ldots, e_n \in E \setminus \mathcal{X} \text{ and } \sigma = mgu(\left\{e_i \stackrel{?}{=} l_i\right\}_{1\le i\le n}) \\ \mathcal{C}'_\beta \textit{ is obtained from } \mathcal{C}_\beta \textit{ by} \\ \textit{adding } r \textit{ to left hand side of constraints} \end{array}$$

Fig. 3. System of transformation rules.

The progress proof relies on two normalisation lemmas for constraint systems.

Lemma 6. *Let $\mathcal{C} = (\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta)$ be a constraint system such that $\mathcal{C}_\alpha$ is in solved form. Then, for all substitutions σ we have: $\sigma \models \mathcal{C}$ if and only if $\sigma \models (\mathcal{C}_\alpha, (E \setminus \mathcal{X}) \rhd t, \mathcal{C}_\beta)$.*

Proof. It suffices to prove that if $x \in E \cap \mathcal{X}$ and σ is a substitution such that $\sigma \models \mathcal{C}$, then we have $\sigma \models (\mathcal{C}_\alpha, (E \setminus \{x\}) \rhd t, \mathcal{C}_\beta)$. Given $x \in E$, by definition 6, there exists a set of terms $E_x \subseteq E$ such that $E_x \rhd x \in \mathcal{C}_\alpha$. Since $\sigma \models \mathcal{C}$ we have $\sigma \models E_x \rhd x$, and by the fact that $E_x \subseteq E \setminus \{x\}$ we have $\sigma \models E \setminus \{x\} \rhd x$. Since we also have $\sigma \models (E \rhd t)$ then, $\sigma \models E \setminus \{x\} \rhd t$. The reciprocal is obvious since $E \setminus \mathcal{X} \subseteq E$. □

Lemma 7. *Let $\mathcal{C} = (\mathcal{C}_\alpha, E \rhd x, \mathcal{C}_\beta)$ be a constraint system such that $\mathcal{C}_\alpha$ is in solved form and $x \notin \mathrm{Var}(\mathcal{C}_\alpha, E, \mathcal{C}_\beta)$ and let $\mathcal{C}' = (\mathcal{C}_\alpha, \mathcal{C}_\beta)$. We have:*

1. *If $\sigma \models \mathcal{C}$ then $\sigma \models \mathcal{C}'$.*
2. *If $\sigma' \models \mathcal{C}'$ then we can extend σ' to σ such that $\sigma \models \mathcal{C}$.*

Proof. 1. Let $\mathcal{C} = (\mathcal{C}_\alpha, E \rhd x, \mathcal{C}_\beta)$ and let σ be a closed substitution such that $\sigma \models \mathcal{C}$. Since $x \notin \mathrm{Var}(\mathcal{C}_\alpha, E, \mathcal{C}_\beta)$, we deduce that $\mathcal{C}' = (\mathcal{C}_\alpha, \mathcal{C}_\beta)$ is deterministic and $\sigma \models \mathcal{C}'$.

2. Let σ' be a closed substitution such that $\sigma' \models \mathcal{C}'$. Since $\mathrm{Var}(E) \subseteq \mathrm{Var}(\mathcal{C}_\alpha)$, σ' is defined on $\mathrm{Var}(\mathcal{C}_\alpha, E, \mathcal{C}_\beta)$ and since $x \notin \mathrm{Var}(\mathcal{C}_\alpha, \mathcal{C}_\beta)$, $\sigma'(x)$ is not defined. We extend σ' to σ as follows:
   $\sigma(y) = \sigma'(y)$ for $y \in \mathrm{Supp}(\sigma')$, $\sigma(x)$ is a closed term in E.
   Since $x \notin \mathrm{Var}(\mathcal{C}_\alpha, \mathcal{C}_\beta, E)$ and $x\sigma \in E\sigma$, we deduce that $\sigma \models \mathcal{C}$.

□

Simplification step. Let $\mathcal{C} = (\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta)$ be a constraint system such that $\mathcal{C}_\alpha$ in solved form and $t \notin \mathcal{X}$. If we apply *Reduce 1* (resp. *Reduce 2*) on $\mathcal{C}$ using a rule $l_x, l_1, \ldots, l_n \twoheadrightarrow r$ such that there is a variable $x \in l_x \setminus \mathrm{Var}(l_1, \ldots, l_n, r)$ then the constraint $E \rhd x$ will be in the obtained constraint system $\mathcal{C}'$ and x does not appear twice in $\mathcal{C}'$. By lemma 7, this constraint can be deleted from $\mathcal{C}'$. As a consequence, we apply a simplification step on the saturated deduction system $\mathcal{L}'$ that eliminates variables $x \in l_x \setminus \mathrm{Var}(l_1, \ldots, l_n, r)$ for all rules $l_x, l_1, \ldots, l_n \twoheadrightarrow r \in \mathcal{L}'$.

Each of the rules in Fig. 3 is correct and complete w.r.t. the satisfiability of constraint systems.

Lemma 8. *A satisfiable constraint system not in solved form can be reduced into another satisfiable constraint system by applying a rule of figure 3.*

Proof. Let $\mathcal{C} = (E_j \rhd t_j)_{1 \leq j \leq n}$ be a satisfiable constraint system not in solved form and let i be the smallest integer such that $t_i \notin \mathcal{X}$. Let $\mathcal{C} = (\mathcal{C}_\alpha, E_i \rhd t_i, \mathcal{C}_\beta)$ where $\mathcal{C}_\alpha$ is in solved form. Since $\mathcal{C}$ is satisfiable there exists a substitution σ such that $\sigma \models_{\mathcal{I}'} \mathcal{C}$. Let us prove that $\mathcal{C}$ can be reduced into another satisfiable constraint system $\mathcal{C}'$ by applying transformation rules given in figure 3. By lemma 6, $\sigma \models_{\mathcal{I}'} \mathcal{C}$ implies $\sigma \models_{\mathcal{I}'} (\mathcal{C}_\alpha, E_i \setminus \mathcal{X} \rhd t_i, \mathcal{C}_\beta)$ and that, by lemma 3, there is a well-formed derivation D starting from $(E_i \setminus \mathcal{X})\sigma$ of goal $t_i\sigma$. We have two cases:

– If $t_i\sigma \in (E_i \setminus \mathcal{X})\sigma$ then there exists a term $u \in E_i \setminus \mathcal{X}$ such that $u\sigma = t_i\sigma$. Let $\mu = mgu(t_i, u)$, we have $\sigma = \mu\theta$ for some substitution θ. $\mathcal{C}$ can then be reduced to $\mathcal{C}'$ by applying *Unif* rule, $\mathcal{C}' = (\mathcal{C}_\alpha\mu, \mathcal{C}_\beta\mu)$ and $\theta \models_{\mathcal{I}'} \mathcal{C}'$.
– If $t_i\sigma \notin (E_i \setminus \mathcal{X})\sigma$, let $D : (E_i \setminus \mathcal{X})\sigma \twoheadrightarrow \ldots \twoheadrightarrow F\sigma, t_i\sigma$ and for every step in D where $l \twoheadrightarrow r$ is the rule applied with the substitution γ, for every $s \in l \setminus \mathcal{X}$, we have either $s\gamma \in (E_i \setminus \mathcal{X})\sigma$ or $s\gamma$ was constructed by a former decreasing rule.
  - Suppose that all applied rules in D are increasing and let $l \twoheadrightarrow r$ be the last applied rule with the substitution γ, this implies that $r\gamma = t_i\sigma$ and for every $s \in l \setminus \mathcal{X}$, $s\gamma \in (E_i \setminus \mathcal{X})\sigma$ and then for every $s \in l \setminus \mathcal{X}$ there exists a term $u \in E_i \setminus \mathcal{X}$ such that $s\gamma = u\sigma$. Let μ be the most general unifier of $\left\{r \stackrel{?}{=} t_i, (s \stackrel{?}{=} u)_{\forall s \in l\setminus\mathcal{X}, u \in E_i\setminus\mathcal{X} \text{ and } s\gamma = u\sigma}\right\}$, we have $\sigma = \mu\theta$ and $\gamma = \mu\theta$ for some θ. This implies that $\mathcal{C}$ can be reduced to $\mathcal{C}' = (\mathcal{C}_\alpha, (E_i \rhd x)_{x\in l}, \mathcal{C}_\beta)\mu$ by applying *Reduce 1* and $\theta \models_{\mathcal{I}'} \mathcal{C}'$.
  - Suppose that D contains decreasing rules and let j be the first step where the applied rule is decreasing. Let $l \twoheadrightarrow r$ be this rule applied with substitution γ. $D : (E_i \setminus \mathcal{X})\sigma = F_0\sigma \twoheadrightarrow F_0\sigma, t_1\sigma \twoheadrightarrow \ldots \twoheadrightarrow F_{j-1}\sigma \twoheadrightarrow F_{j-1}\sigma, t_j\sigma \twoheadrightarrow \ldots \twoheadrightarrow F_{n-1}\sigma, t_i\sigma$. Since D is well-formed, we deduce that for every $s \in l \setminus \mathcal{X}$, $s\gamma \in (E_i \setminus \mathcal{X})\sigma$ and then, for every $s \in l \setminus \mathcal{X}$ there exists a term $u \in E_i \setminus \mathcal{X}$ such that $s\gamma = u\sigma$. Let μ be the most general unifier, we have $\gamma = \mu\theta$ and $\gamma = \mu\theta$ for some substitution θ. This implies that $\mathcal{C}$ can be reduced to $\mathcal{C}' = (\mathcal{C}_\alpha, (E_i \rhd x)_{x\in l}, (E_i \cup r) \rhd t_i, \mathcal{C}'_\beta)\mu$ by applying *Reduce 2* and $\theta \models_{\mathcal{I}'} \mathcal{C}'$.

□

Lemma 9. *Let $\mathcal{C}$ and $\mathcal{C}'$ be two constraint systems such that $\mathcal{C}'$ is obtained from $\mathcal{C}$ by applying a transformation rule. If $\mathcal{C}$' is satisfiable then so is $\mathcal{C}$.*

Proof. Let $\mathcal{C}$ and $\mathcal{C}'$ be two constraint systems such that $\mathcal{C}'$ is obtained from $\mathcal{C}$ by applying a transformation rule and suppose that $\mathcal{C}'$ is satisfiable. Let σ' be a solution of $\mathcal{C}'$ and let us prove that $\mathcal{C}$ is satisfiable. Since a transformation rule can be applied on $\mathcal{C}$, $\mathcal{C}$ can't be in solved form. Suppose that $\mathcal{C} = (\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta)$ where $\mathcal{C}_\alpha$ is in solved form and $t \notin \mathcal{X}$.

– If $\mathcal{C}'$ is obtained from $\mathcal{C}$ by applying *Unif* rule then, there exists a term $u \in E \setminus \mathcal{X}$ such that u and t are unifiable. Let μ be the most general unifier then $\mathcal{C}' = (\mathcal{C}_\alpha\mu, \mathcal{C}_\beta\mu)$. Since $\sigma' \models_{\mathcal{I}'} \mathcal{C}'$, we have $\sigma' \circ \mu \models_{\mathcal{I}'} (\mathcal{C}_\alpha, \mathcal{C}_\beta)$ and by the fact that μ is the most general unifier of t and a term in E we have $\sigma' \circ \mu \models_{\mathcal{I}'} E \rhd t$. We deduce that $\sigma' \circ \mu \models_{\mathcal{I}'} \mathcal{C}$.
– If $\mathcal{C}'$ is obtained from $\mathcal{C}$ by applying *Reduce 1* then there exists an increasing rule $l_x, l_1, \ldots, l_n \twoheadrightarrow r$, a set of terms $e_1, \ldots, e_n$ in $E \setminus \mathcal{X}$ such that $\left\{r \stackrel{?}{=} t, (l_i \stackrel{?}{=} e_i)_{1\le i\le n}\right\}$ has solution. Let μ be the most general unifier. $\mathcal{C}' = (\mathcal{C}_\alpha, (E \rhd x)_{x\in l_x}, \mathcal{C}_\beta)\mu$. Since $\sigma' \models_{\mathcal{I}'} \mathcal{C}'$ and by definition of μ, we have $\sigma' \circ \mu \models_{\mathcal{I}'} \mathcal{C}$.

– If $\mathcal{C}'$ is obtained from $\mathcal{C}$ by applying *Reduce 2* then there exists a decreasing rule $l_x, l_1, \ldots, l_n \twoheadrightarrow r$ and a set of terms $e_1, \ldots, e_n$ in $E \setminus \mathcal{X}$ such that $\left\{(l_i \stackrel{?}{=} e_i)_{1 \leq i \leq n}\right\}$ has solution. Let μ be the most general unifier. $\mathcal{C}' = (\mathcal{C}_\alpha, (E \rhd x)_{x \in l_x}, (E \cup r) \rhd t, \mathcal{C}'_\beta)\mu$. Since $\sigma' \models_{\mathcal{I}'} \mathcal{C}'$ and by definition of μ and constraint systems, we have $\sigma' \circ \mu \models_{\mathcal{I}'} \mathcal{C}$.

□

5 Decidability of reachability problems

In this section we first prove that if the saturation terminates then ground reachability problems are decidable. We then give an additional criterion that will permit us to lift this result to general reachability problems.

5.1 Decidability of ground reachability problems

We recall that $\mathcal{I}_0 = \langle \mathcal{G}, \mathcal{L}_0, \mathcal{H} \rangle$ is the initial deduction system and $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$ is the saturated deduction system.

Let us also first recall in the following lemma some properties of reduction ordering.

Lemma 10. *Let $t_1, t_2 \in \mathrm{T}(\mathcal{G}, \mathcal{X})$ and $t_1 \preceq t_2$. We have:*

1. $\mathrm{Var}(t_1) \subseteq \mathrm{Var}(t_2)$
2. $t_2 \notin \mathrm{SSub}(t_1)$
3. *If $t_2 \in \mathcal{X}$ then $t_1 = t_2$*
4. *If $t_1 \notin \mathcal{X}$ then $t_1 \not\prec x$*

Proof. 1. Let t_1 and t_2 be two terms and $t_1 \preceq t_2$. If $t_1 = t_2$ then we have obviously $\mathrm{Var}(t_1) = \mathrm{Var}(t_2)$. Suppose $t_1 \neq t_2$ this implies that $t_1 \prec t_2$ and let us prove that $\mathrm{Var}(t_1) \subseteq \mathrm{Var}(t_2)$. By contradiction, suppose that $\mathrm{Var}(t_1) \not\subseteq \mathrm{Var}(t_2)$ and let $x \in \mathrm{Var}(t_1) \setminus \mathrm{Var}(t_2)$. By definition of $\prec$, we have $t_1\sigma \prec t_2\sigma$ for all substitutions σ. Let σ be a substitution such that $\mathrm{Supp}(\sigma) = \{x\}$ and $\sigma(x) = t_2$. This implies that $t_2\sigma = t_2$ and $t_2 \in \mathrm{Sub}(t_1\sigma)$ which contradicts $t_1 \prec t_2$.

2. If $t_2 \in \mathrm{SSub}(t_1)$ this implies that $t_1 \neq t_2$ and $t_2 \prec t_1$ which contradicts $t_1 \preceq t_2$.
3. If $t_2 = x$ we deduce that $\mathrm{Var}(t_1) \subseteq \{x\}$ and $x \notin \mathrm{SSub}(t_1)$. This implies that $t_1 = x$.
4. Suppose that $t_1 \neq x$ and $t_1 \prec x$. This implies that $\mathrm{Var}(t_1) \subseteq \{x\}$ and then, either $t_1 = x$ or $x \in \mathrm{SSub}(t_1)$. This contradicts the fact that $t_1 \neq x$ and $x \notin \mathrm{SSub}(t_1)$.

□

A core result of this paper is the following lemma.

Lemma 11. *Let $\mathcal{I}'$ be a saturated deduction system such that $\mathcal{L}'$ is finite. Applying the transformation algorithm of Fig. 3 on a constraint system $\mathcal{C}$ without instantiating the variables of $\mathcal{C}$ yields only a finite number of different constraint systems.*

Proof. Assume the application of rules of Fig. 3 yields an infinite sequence of constraint systems $\mathcal{C}_1, \ldots, \mathcal{C}_n, \ldots$. Let us prove there is only a finite number of different $\mathcal{C}_i$ when identical constraints within a constraint system are identified.

Let us first prove that there is only a finite number of different left-hand side of deduction constraints. The number of different left-hand sides in a constraint system does not change (or decrease) when a Unif or Reduce1 rule is applied. Assume now that a decreasing rule $l_x, l_1, \ldots, l_n \twoheadrightarrow r \in \mathcal{L}'$ is applied with a substitution σ on a constraint with left-hand side E. If $r\sigma \in E$, the number of different left-hand side does not change. Thus let us assume $r\sigma \notin E$, and thus $r\sigma \notin \cup\{l_1\sigma, \ldots, l_n\sigma\}$. Since r is smaller or equal to a term of the left-hand side of there rule, we have two case:

- Either there exists i with $l_i\sigma \succ r\sigma$, and thus there exists $e \in E$ such that $e \succ r\sigma$.
- Or $r \in l_x \setminus \mathrm{Var}(l_1, \ldots, l_n)$. Then the obtained constraint system contains the deduction constraints $E \rhd r$ and $E \cup \{r\} \rhd t$ and not other constraint contains r. By Lemma 6 the obtained constraint system is equivalent to the one in which $E \cup \{r\} \rhd t$ is replaced by $E \rhd t$.

Let us now consider the set T which is the union of all left-hand side of deduction constraints reachable from E by employing a decreasing rule.

- the root is labelled by $\emptyset$;
- the sons of the root are labelled by the terms in a left-hand side E;
- The sons of the non-root node are defined as follows: assume there exists two left-hand sides E' and E'' where E' is reachable from E, and there is a decreasing rule whose application leads to the addition of a deduction constraint with left-hand side $E'' = E'', t_1$. Let $t_2 \in E'$ be the term strictly greater than t_1. We then set t_1 as a son of t_2.

Since $t_2 \succ t_1$ there is no cycle, and since we consider sets reachable from E, the "is son of" relation is connected. It thus defines a tree. We note that t_2 is the instance of a non-variable term l in the left-hand side of a decreasing rule. There is only a finite number of such terms. Since we consider deductions in the empty theory, for each l there is a unique substitution σ such that $l\sigma = t_2$. Given the above properties of reduction ordering we have $\mathrm{Var}(r) \subseteq \mathrm{Var}(l)$ and thus $t_1 = r\sigma$ is uniquely determined by the rule applied. Thus, each term t_2 has a finite number of sons t_1. Along each branch of the tree a node t is strictly smaller than its parent. Since $\succ$ is a well-founded ordering, this implies that each branch is finite. Thus, by König's Lemma, this tree is finite. We conclude that T itself is finite. Each left-hand side of a deduction constraint is a subset of T, thus there is only a finite number of different left-hand sides.

When applying REDUCE 1 or REDUCE 2 on a constraint $E \rhd t$, the newly introduced constraints $E \rhd t'$ are such that t' is a strict subterm of a term in E or t. Let $E' \rhd t'$ be a deduction constraint reached from $E \rhd t$. Either t' is a subterm of t or there exists E'' reachable from E such that t' is a strict subterm of E''. Since there is only a finite number of different E'', there is thus only a finite number of possible right-hand side of constraints.

In conclusion only a finite number of deduction constraints $E' \rhd t'$ can be reached from a deduction constraint $E \rhd t$. Thus only a finite number of constraint systems can be reached from a given one by applying rules that do not instantiate the variables in the constraint system. □

Definition 7. *An $\mathcal{I}_0$-ground constraint system $\mathcal{C}$ is denoted $(E_1 \rhd t_1, \ldots, E_n \rhd t_n)$ and is defined by a sequence of pairs $(E_i, t_i)_{i \in \{1,\ldots,n\}}$ such that E_i (resp. t_i) is a set of ground terms (resp. ground term) in normal form and $E_i \subseteq E_{i+1}$ for $i \in \{1, \ldots, n\}$.*

We note that an $\mathcal{I}_0$-ground constraint $E \rhd t$ is valid if $t \in \overline{E}^{\mathcal{I}_0}$. We now consider the following problem:

$\mathcal{I}_0$-Ground Reachability Problem

Input: An $\mathcal{I}_0$-ground constraint system $\mathcal{C}$.

Output: VAL iff $(t_i \in \overline{E_i}^{\mathcal{I}_0})_{i \in \{1,\ldots,n\}}$.

We recall that $t \in \overline{E}^{\mathcal{I}_0}$ iff $t \in \overline{E}^{\mathcal{I}'}$ while E (resp. t) is set of closed terms (resp. closed term) in normal form (Lemmas 1 and 2). This implies that solving $\mathcal{I}_0$-ground reachability problem is reduced to solving $\mathcal{I}'$-ground reachability problem. It is then routine to see that a ground constraint system is valid if, and only if, it reduces to an empty sequence of deduction constraints. Thus by Lemma 11 we have:

Theorem 1 *If the saturation algorithm terminates on $\mathcal{L}_0$, the $\mathcal{I}_0$-ground reachability problem is decidable.*

6 Termination of Saturation does not imply decidability of general reachability problems

It is well-known how to encode 2-stack automata into deduction systems. However the saturation will typically not terminate on standard encodings as it will amount in this case to the pre-computation of all possible executions of the automaton. We can however adapt the construction so that saturation terminates. We consider a signature $\mathcal{G}$ such that, for all symbol $f \in \mathcal{G}_0$ of arity n, there is a deduction rule $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n)$, and the signature $\mathcal{G} = \mathcal{G} \cup \{g\}$ with g a symbol of arity 1. Let $(Q, Q_I, Q_F, \Sigma, \Pi, \Delta)$ be a finite 2-stack automaton, where Q is the finite set of states of the automaton, Q_I and Q_F its initial and final states, Σ denotes the alphabet of the words read by the automaton, and Π

denotes the elements in the stacks of the automaton. We shall encode the emptiness of the language recognised by this automaton into a general reachability problem. Let us assume there exists:

- $\bot \in \mathcal{G}_0$ be a constant denoting the empty stack or the empty word;
- one unary symbol u_α for each letter $\alpha \in \Sigma \cup \Pi$;
- one constant $q \in \mathcal{G}$ for each state in Q;
- one symbol $s \in \mathcal{G}$ of arity 4 where we intend that:
  - the first argument represents the word that remains to be read by the automaton;
  - the second argument represents the current state of the automaton;
  - the third and fourth arguments represent the two stacks of the automaton.
- one symbol f of arity 2.

We represent a transition from a state σ_1 to a state σ_2 with a symbol τ of arity 1 and a rewriting rule $\tau(g(f(\sigma_1, f(\sigma_2, x)))) \to g(f(\sigma_2, x))$. The rewriting system has no critical pairs, and thus is confluent. Since every narrowing step decreases strictly the number of "τ" symbols in a term, narrowing terminated, and thus the equational theory has the finite variant property. At the end of the first step of the saturation the system will contain the rules enabling the attacker to build sequences of states, and additional rules $g(f(\sigma_1, f(\sigma_2, x))) \twoheadrightarrow g(f(\sigma_2, x))$ that are decreasing for any recursive path ordering. Since there is no increasing rule with the symbol g in the right-hand side, we leave to the reader the proof that saturation terminates, and hence that ground reachability problems are decidable.

However, the instance of x in the following reachability problem encodes a word recognised by the automaton after a run encoded by the instance of y:

$$\emptyset \rhd f(s(x, q_0, \bot, \bot), y), g(f(s(x, q_0, \bot, \bot), y)) \rhd g(s(\bot, q_f, \bot, \bot))$$

This example proves (with $q_0 \in Q_I$ and $q_f \in Q_F$) that the saturation can terminate and yield a deduction system for which general reachability problems are not decidable.

The undecidability comes from the fact that one can apply an unbounded number of decreasing rules on a non-ground terms, and from the "lack of regularity" on the terms obtained.

7 Decidability of general reachability problems

We recall that the initial intruder system is given by $\mathcal{I}_0 = \langle \mathcal{G}, \mathcal{L}_0, \mathcal{H} \rangle$ while $\mathcal{H}$ is generated by a convergent equational theory and has the finite variant property. We recall also that $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$ is the saturated intruder system.

We give here a simple criterion that permits to ensure the termination of the resolution of a constraint problem with a saturated deduction system. Let T be a set of terms, $T = \{t_1, \ldots, t_m\}$, we let $\Delta(T)$ to be the set of strict maximal subterms of T and we define:

$$\delta(T) = \begin{cases} +\infty & \text{if } T \subseteq \mathcal{X} \\ |T \setminus \mathcal{X}| - |\mathrm{Var}(T \setminus \mathcal{X}) \setminus (T \cap \mathcal{X})| & \text{otherwise.} \end{cases}$$

Now let us define $\mu(T)$. We consider the image of the set of terms T by the rewriting system $\mathcal{U}$ containing rules $f(x_1, \ldots, x_n) \to x_1, \ldots, x_n$ for every symbol f in the signature of the deduction system. We define:

$$\mu(T) = \min_{\substack{T\sigma \to_{\mathcal{U}}^* T' \\ \sigma \text{ mgu of subterms of } T}} \delta(T')$$

We extend μ to rules as follows. Let $\mathcal{L}'$ be the set of deduction rules. We recall that $\mathcal{L}'$ is partitioned into two disjoint sets of deduction rules, the set of increasing rules $\mathcal{L}'_{inc}$ and the set of decreasing rules $\mathcal{L}'_{dec}$. For every rule $l \twoheadrightarrow r \in \mathcal{L}'$,

$$\mu(l \twoheadrightarrow r) = \begin{cases} \mu(\Delta(l \setminus \mathcal{X} \cup \{r\})) \text{ if } l \twoheadrightarrow r \text{ is increasing,} \\ \mu(\Delta(l \setminus \mathcal{X})) \text{ otherwise.} \end{cases}$$

Definition 8. *(Contracting deduction systems) A saturated deduction system* $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$ *is* contracting *if for all rules* $l \twoheadrightarrow r$ *in* $\mathcal{L}'$ *we have* $\mu(l \twoheadrightarrow r) > 0$.

Lemma 12. *Let* $S = \{s_1, \ldots, s_n\}$ *and* $T = \{t_1, \ldots, t_n\}$ *be two sets of terms and let* σ *be the most general unifier of* $V = \left\{s_1 \stackrel{?}{=} t_1, \ldots, s_n \stackrel{?}{=} t_n\right\}$. *If* $\mu(T) > 0$ *then either* $|\mathrm{Var}(s_1, \ldots, s_n)| > |\mathrm{Var}((s_1, \ldots, s_n, t_1, \ldots, t_n)\sigma)|$ *or* $|\mathrm{Var}(s_1, \ldots, s_n)| = |\mathrm{Var}((s_1, \ldots, s_n, t_1, \ldots, t_n)\sigma)|$, $S = S\sigma$ *and for all* $x \in \mathrm{Var}(T)$ *there is* $i \in \{1, \ldots, n\}$ *such that* $\sigma(x) \preceq s_i$.

Proof. Let $V = \left\{s_1 \stackrel{?}{=} t_1, \ldots, s_n \stackrel{?}{=} t_n\right\}$. In order to solve V, we apply the first step of the unification algorithm of Martelli-Montanari [27]. We reduce V to $V' = \left\{x_1 \stackrel{?}{=} u_1, \ldots, x_k \stackrel{?}{=} u_k, x_{k+1} \stackrel{?}{=} u_{k+1}, \ldots, x_m \stackrel{?}{=} u_m\right\}$ such that for every equation $x \stackrel{?}{=} u \in V'$, we have either $x \in \mathrm{Var}(S)$ and $u \in \mathrm{Sub}(T)$ or $x \in \mathrm{Var}(T)$ and $u \in \mathrm{Sub}(S)$. We suppose that $x_j \in \mathrm{Var}(T)$ for $j \in \{1, \ldots, k\}$.

- If $k = m$ then we have $x_j \in \mathrm{Var}(T)$ for $j \in \{1, \ldots, m\}$. We suppose that $x_i \neq x_j$ for all $i, j \in \{1, \ldots, m\}$ and $i \neq j$. This implies that $S\sigma = S$ and $\mathrm{Var}(T)$ are instantiated by subterms of S, that is $\mathrm{Var}(T)\sigma$ are smaller or equal than terms in S. We conclude also that $|\mathrm{Var}(s_1, \ldots, s_n)| = |\mathrm{Var}((s_1, \ldots, s_n, t_1, \ldots, t_n)\sigma)|$.
- If $k \neq m$ assume $\{u_{k+1}, \ldots, u_m\} \notin \mathrm{Var}(T)$, we have different cases:
  - If for all different $i, j \in \{1, \ldots, m\}$ we have $x_i \neq x_j$ then $m - k$ variables of S, $x_{k+1}, \ldots, x_m$, are instantiated by subterms of T, $u_{k+1}, \ldots, u_m$. This implies that when we apply σ to $\mathcal{S}$, new variables, $\mathrm{Var}(u_{k+1}, \ldots, u_m) \setminus \{x_1, \ldots, x_k\}$ will appear in $S\sigma$. There exists a set $T' \not\subseteq \mathcal{X}$ such that $T \to_{\mathcal{U}}^* T'$ and $T' = \{x_1, \ldots, x_k, u_{k+1}, \ldots, u_m\}$. Since $\mu(T) > 0$, we have $|T' \setminus \mathcal{X}| > |\mathrm{Var}(T' \setminus \mathcal{X}) \setminus (x_1, \ldots, x_k)|$. This implies that $|\mathrm{Var}(s_1, \ldots, s_n)| > |\mathrm{Var}((s_1, \ldots, s_n, t_1, \ldots, t_n)\sigma)|$.

- If there is different $i, j \in \{1, \ldots, m\}$ such that $x_i = x_j$:
  - If $i, j \leq k$ then we have to unify two subterms of S. Let u_i and u_j be these two subterms and α be their most general unifier.
    Let us apply α on V and to solve V we have to solve $V\alpha = \left\{s_1\alpha \stackrel{?}{=} t_1, \ldots, s_n\alpha \stackrel{?}{=} t_n\right\}$. To solve $V\alpha$ we reduce it to another system V" where equations have the same form as in V'. We note that $|\text{Var}(T)|$ in $V\alpha$ is the same as in V and $|\text{Var}(S)|$ is reduced.
    By the same reasoning as above, we deduce that $|\text{Var}(S)| > |\text{Var}(S\sigma, T\sigma)|$.
  - If $i, j > k$ then we have to unify two subterms of T. Let u_i and u_j be these two subterms and α be their most general unifier. Let us apply α on V and to solve V we have to solve $V\alpha = \left\{s_1 \stackrel{?}{=} t_1\alpha, \ldots, s_n \stackrel{?}{=} t_n\alpha\right\}$ and to solve $V\alpha$, we have to reduce it to another system V" where equations have the same form as in V'. V" $= \left\{x_1 \stackrel{?}{=} u_1, \ldots, x_m \stackrel{?}{=} u_m\right\}$ where $x_1 \ldots, x_k \in \text{Var}(T\alpha)$ and $x_{k+1}, \ldots, x_m \in \text{Var}(S)$. By definition of μ and by following the same reasoning as above, we deduce that:
    - If $k = m$ and for all different $i, j \in \{1, \ldots, m\}$ we have $x_i \neq x_j$, we deduce that $S = S\sigma$, $\text{Var}(T)\sigma$ are smaller or equals than terms in S and then $|\text{Var}(S)| = |\text{Var}(S\sigma, T\sigma)|$.
    - If $k = m$ and there is different i, j such that $x_i = x_j$ then we have to unify two subterms of S and then we conclude that $|\text{Var}(S)| > |\text{Var}(S\sigma, T\sigma)|$.
    - If $k \neq m$ we deduce that $|\text{Var}(S)| > |\text{Var}(S\sigma, T\sigma)|$.

□

The definition of μ is tailored to the proof of the following Lemma.

Remark. Let T be a set of terms and let $\Sigma(T) = \{\sigma \textit{ s.t. } \sigma \textit{ is the most general unifier of some subterms of } T\}$. We remark that $\mu(T)$ is defined with respect to $T\sigma$ for every $\sigma \in \Sigma$. It will be more naturel and more general if $\mu(T)$ is defined with respect to T instead of some instances of T. The so-called general definition will be defined as follow:

$$\mu(T) = \min_{T \to_{\mathcal{U}}^{*} T'} \delta(T')$$

Using the general definition of μ, we remark that $\mu(T) > 0$ does not imply $\mu(T\sigma) > 0$ for a set of terms T and a subtitution $\sigma \in \Sigma(T)$. Let $T = \{f(x, x), f(x, y), f(y, x)\}$ and let σ be such that $\sigma(x) = y$. Using the general definition of μ, we remark that $\mu(T) > 0$ and $\mu(T\sigma) = 0$.

Unfortunately, the lemma 12, used in the proof of termination (lemma 13), becomes false with the general definition.

Lemma 13. *Let $\mathcal{I}'$ be a saturated contracting deduction system, $\mathcal{C}$ be a $\mathcal{I}'$-constraint system not in solved form. If a transformation is applied on $\mathcal{C}$ to yield a constraint system $\mathcal{C}'$, then either the substitution applied does not instantiate the variables of $\mathcal{C}$ and $\mathrm{Var}(\mathcal{C}') \subseteq \mathrm{Var}(\mathcal{C})$ or $|\mathrm{Var}(\mathcal{C}')| < |\mathrm{Var}(\mathcal{C})|$.*

Proof. Let $\mathcal{C}$ be a constraint system such that a transformation rule can be applied on it. This implies that $\mathcal{C}$ is not in solved form. Let $\mathcal{C} = (\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta)$ such that $\mathcal{C}_\alpha$ is in solved form and $t \notin \mathcal{X}$. We have three cases:

- If we apply *Unif* rule on $\mathcal{C}$ then there exists a term $e \in E \setminus \mathcal{X}$ such that e and t are unifiable and σ is the most general unifier. $\mathcal{C}$ is then reduced to $\mathcal{C}' = (\mathcal{C}_\alpha, \mathcal{C}_\beta)\sigma$. Since we unify two subterms of $\mathcal{C}$ in the empty theory, either σ does not instantiate the variables of $\mathcal{C}$ and then $\mathcal{C}' = (\mathcal{C}_\alpha, \mathcal{C}_\beta)$ (which implies that $\mathrm{Var}(\mathcal{C}') \subseteq \mathrm{Var}(\mathcal{C})$) or σ instantiates the variables of $\mathcal{C}$ (and then $|\mathrm{Var}(\mathcal{C}')| < |\mathrm{Var}(\mathcal{C})|$).
- Assume we apply Reduce 1 on $\mathcal{C}$. By definition of Reduce 1 there exists an increasing rule $l_x, l_1, \ldots, l_n \twoheadrightarrow r \in \mathcal{L}'$, a set of terms $e_1, \ldots, e_n \in E \setminus \mathcal{X}$ such that $\mathcal{S} = \left\{ r \stackrel{?}{=} t, (e_i \stackrel{?}{=} l_i)_{1 \leq i \leq n} \right\}$ has a solution. Let σ be its most general unifier. Either $\sigma_{|\mathrm{Var}(\mathcal{C})} = Id[\sigma_{|\mathrm{Var}(\mathcal{C})} = Id]$ or not. Let us examine the two cases.

  $\sigma_{|\mathrm{Var}(\mathcal{C})} = Id$**.** In this case, $\mathcal{C}$ is reduced to $\mathcal{C}' = (\mathcal{C}_\alpha, (E \rhd x\sigma)_{x \in l_x}, \mathcal{C}_\beta)$. For each $l_i \in \{l_1, \ldots, l_n\}$ we have, by definition of σ, $l_i\sigma = e_i$. Also, we have $r\sigma = t$. Thus for each $x \in \mathrm{Var}(l_1, \ldots, l_n, r)$ we have $\mathrm{Var}(x\sigma) \subseteq \mathrm{Var}(\mathcal{C})$. Since $l_x \subseteq \mathrm{Var}(l_1, \ldots, l_n, r)$ we deduce that $\mathrm{Var}(\mathcal{C}') \subseteq \mathrm{Var}(\mathcal{C})$.

  $\sigma_{|\mathrm{Var}(\mathcal{C})} \neq Id$**.** In this case $\mathcal{C}$ is reduced to $\mathcal{C}' = (\mathcal{C}_\alpha, (E \rhd x)_{x \in l_x}, \mathcal{C}_\beta)\sigma$. Since the e_i and r are not variables, we can decompose all equations in $\mathcal{S}$ to obtain a set of equations in which each equation has a member in $\Delta(l_1, \ldots, l_n, r)$. Since the deduction system is contracting Lemma 12 implies $|\mathrm{Var}(e_1, \ldots, e_n, t)| > |\mathrm{Var}(e_1\sigma, \ldots, e_n\sigma, t\sigma, l_1\sigma, \ldots, l_n\sigma, r\sigma)|$. Since $l_x \subseteq \mathrm{Var}(l_1, \ldots, l_n, r)$ we deduce that $|\mathrm{Var}(\mathcal{C})| > |\mathrm{Var}(\mathcal{C}')|$.
- Let us finally assume Reduce 2 is applied. First let us prove we can assume $l_x \cup \mathrm{Var}(r) \subseteq \mathrm{Var}(\{l_1, \ldots, l_n\})$. Since the rule is decreasing there exists a term $l \in l_x \cup \{l_1, \ldots, l_n\}$ such that $\mathrm{Var}(r) \subseteq \mathrm{Var}(l)$. Thus it suffices to prove $l_x \subseteq \mathrm{Var}(\{l_1, \ldots, l_n\})$. By definition of the Reduce 2 rule, the constraint system $\mathcal{C}$ is transformed into

$$\begin{aligned} & (\mathcal{C}_\alpha, (E \rhd y)_{y \in l_x \setminus \{x\}}, E \rhd x, E \cup \{x\} \rhd t, \mathcal{C}'_\beta)\sigma \\ = \; & \mathcal{C}_\alpha\sigma, (E\sigma \rhd y\sigma)_{y \in l_x \setminus \{x\}}, E\sigma \rhd x, E\sigma \cup \{x\} \rhd t\sigma, \mathcal{C}'_\beta\sigma \\ \equiv \; & \mathcal{C}_\alpha\sigma, (E\sigma \rhd y\sigma)_{y \in l_x \setminus \{x\}}, E\sigma \rhd x, E\sigma \rhd t\sigma, \mathcal{C}_\beta\sigma \\ \equiv \; & \mathcal{C}_\alpha\sigma, (E\sigma \rhd y\sigma)_{y \in l_x \setminus \{x\}}, E\sigma \rhd t\sigma, \mathcal{C}_\beta\sigma \end{aligned}$$

  where the first $\equiv$ is by Lemma 6, and the second one by Lemma 7. Thus the resulting system is equivalent for solutions to one in which $l_x \subseteq \mathrm{Var}(l_1, \ldots, l_n)$. We can then apply the same reasoning as above.

□

We may now conclude by applying the previous results and again König's Lemma.

Theorem 2 *Let $\mathcal{I}_0 = \langle \mathcal{G}, \mathcal{L}_0, \mathcal{H} \rangle$ be a deduction system such that the saturation of $\mathcal{L}_0$ terminates , and the resulting deduction system is contracting. Then the $\mathcal{I}_0$-reachability problem is decidable.*

Proof. It suffices to prove that the application of rules of Fig. 3 terminates. Assume there exists an $\mathcal{I}'$-constraint system $\mathcal{C}$ and an infinite sequence of transformations starting from $\mathcal{C}$. Let $\mathcal{C}_1, \ldots, \mathcal{C}_n, \ldots$ be the resulting sequence of constraint systems. By Lemma 13, at each step $|\mathrm{Var}(\mathcal{C}_i)| \geq |\mathrm{Var}(\mathcal{C}_{i+1})|$ and if there is equality, then the substitution applied on $\mathcal{C}_i$ is the identity (does not instantiate the variables of $\mathcal{C}$). Since we must have a positive number of variables, there is only a finite number of steps where the substitution is not the identity. Let $\mathcal{C}_n$ be the resulting constraint system. Since all subsequent transformation do not instantiate the variables of $\mathcal{C}_n$ and its successor, the sequence has only a finite number of different constraint systems.

Since $\mathcal{L}'$ is finite, each constraint system has only a finite number of successors. Thus by König Lemma there is only a finite number of different constraint systems. □

8 Some relevant equational theories

We give here some examples of well-known equational theories where the saturation applied on the corresponding initial set of deduction rules terminates.

8.1 Dolev-Yao theory with explicit destructors

The Dolev-Yao theory with explicit destructors is the classical Dolev-Yao model with explicit destructors such as decryption and projections. This theory is given by the following set of equations:

$$\mathcal{H}_{DV} = \begin{cases} Dec_s(Enc_s(x,y),y) = x, \\ Enc_s(Dec_s(x,y),y) = x, \\ Dec_a(Enc_a(x,PK(y)),SK(y)) = x, \\ Enc_a(Dec_a(x,SK(y)),PK(y)) = x, \\ \pi_1(\langle x,y\rangle) = x, \\ \pi_2(\langle x,y\rangle) = y. \end{cases}$$

By orienting equations of $\mathcal{H}_{DV}$ from left to right, we obtain a rewrite system $\mathcal{R}_{DV}$ generating $\mathcal{H}_{DV}$. We remark that $\mathcal{R}_{DV}$ is convergent and $\mathcal{H}_{DV}$ has finite variant property.

The initial set of deduction rules is given by the following set of rules:

$$\mathcal{L}_0 = \begin{cases} x, y \twoheadrightarrow \langle x,y\rangle, \\ x \twoheadrightarrow \pi_1(x), \\ x \twoheadrightarrow \pi_2(x), \\ x, y \twoheadrightarrow Enc_a(x,y), \\ x, y \twoheadrightarrow Dec_a(x,y), \\ x, y \twoheadrightarrow Enc_s(x,y), \\ x, y \twoheadrightarrow Dec_s(x,y). \end{cases}$$

The saturatation (modulo the simplification introduced after the lemma 7) outputs the following set of deduction rules:

$$\mathcal{L}' = \mathcal{L}_0 \cup \begin{cases} \langle x, y\rangle \twoheadrightarrow x, \\ \langle x, y\rangle \twoheadrightarrow y, \\ Dec_a(x, SK(y)), PK(y) \twoheadrightarrow x, \\ Enc_a(x, PK(y)), SK(y) \twoheadrightarrow x, \\ Des_s(x, y), y \twoheadrightarrow x, \\ Enc_s(x, y), y \twoheadrightarrow x, \\ x, PK(y), SK(y) \twoheadrightarrow x. \end{cases}$$

8.2 Digital signature theory with duplicate signature key selection property

The theory of digital signature with duplicate signature key selection property is defined in [11] and is given by the following set of equations:

$$\mathcal{H}_{DSKS} = \begin{cases} Ver(x, Sig(x, SK(y)), PK(y)) = 1, \\ Ver(x, Sig(x, SK'(y_1, y_2)), PK'(y_1, y_2)) = 1, \\ Sig(x, SK'(PK(y), Sig(x, SK(y)))) = Sig(x, SK(y)). \end{cases}$$

The equational theory $\mathcal{H}_{DSKS}$ is generated by:

$$\mathcal{R}_{DSKS} = \begin{cases} Ver(x, Sig(x, SK(y)), PK(y)) \to 1, \\ Ver(x, Sig(x, SK'(y_1, y_2)), PK'(y_1, y_2)) \to 1, \\ Ver(x, Sig(x, SK(y)), PK'(PK(y), Sig(x, SK(y)))) \to 1, \\ Sig(x, SK'(PK(y), Sig(x, SK(y)))) \to Sig(x, SK(y)). \end{cases}$$

We remark that $\mathcal{R}_{DSKS}$ is convergent and $\mathcal{H}_{DSKS}$ has the finite variant property.

The initial set of deduction rules is given by the following set of rules:

$$\mathcal{L}_0 = \begin{cases} x, y \twoheadrightarrow Sig(x, y), \\ x, y, z \twoheadrightarrow Ver(x, y, z), \\ x, y \twoheadrightarrow SK'(x, y), \\ x, y \twoheadrightarrow PK'(x, y), \\ \emptyset \twoheadrightarrow 0, \\ \emptyset \twoheadrightarrow 1. \end{cases}$$

The saturatation (modulo the simplification introduced after the lemma 7) outputs the following set of deduction rules:

$$\mathcal{L}' = \mathcal{L}_0 \cup \begin{cases} x, Sig(x, SK(y)), Pk(y) \twoheadrightarrow 1, \\ x, Sig(x, SK'(y_1, y_2)), PK'(y_1, y_2) \twoheadrightarrow 1, \\ x, Sig(x, SK(y)), PK'(PK(y), Sig(x, SK(y))) \twoheadrightarrow 1, \\ x, SK'(PK(y), Sig(x, SK(y))) \twoheadrightarrow Sig(x, SK(y)), \\ SK(y), PK(y) \twoheadrightarrow 1, \\ SK'(y_1, y_2), PK'(y_1, y_2) \twoheadrightarrow 1, \\ x, SK(y), PK'(PK(y), Sig(x, SK(y))) \twoheadrightarrow 1, \\ x, PK(y), Sig(x, SK(y)) \twoheadrightarrow Sig(x, SK(y)), \\ x, PK(y), SK(y) \twoheadrightarrow Sig(x, SK(y)), \\ y_1, y_2, PK'(y_1, y_2) \twoheadrightarrow 1, \\ x, y_1, y_2, Sig(x, SK'(y_1, y_2)) \twoheadrightarrow 1, \\ y_1, y_2, SK'(y_1, y_2) \twoheadrightarrow 1, \\ x, PK(y), Sig(x, SK(y)) \twoheadrightarrow 1, \\ x, PK(y), SK(y), Sig(x, SK(y)) \twoheadrightarrow 1, \\ x, SK(y), PK(y), PK'(PK(y), Sig(x, SK(y))) \twoheadrightarrow 1, \\ x, SK(y), PK(y) \twoheadrightarrow Sig(x, SK(y)). \end{cases}$$

9 Decidability of ground reachability problems for the blind signature theory

Blind signature was introduced in [23], it is defined by the signature $\mathcal{G} = \{Sig, Ver, Bl, Ubl, PK, SK\}$ which satisfies the following set of equations:

$$\mathcal{H} = \begin{cases} Ver(Sig(x, SK(y)), PK(y)) = x, \\ Ubl(Bl(x, y), y) = x, \\ Ubl(Sig(Bl(x, y), SK(z)), y) = Sig(x, SK(z)). \end{cases}$$

Let $\mathcal{R}$ be the set of rules obtained by orienting equations of $\mathcal{H}$ from left to right, $\mathcal{R}$ is convergent and it is obvious that any basic narrowing derivation [22] issuing from any of the right hand side term of the rules of $\mathcal{R}$ terminates. This implies that any narrowing derivation (and in particular basic narrowing derivation) issuing from any term terminates [22] and thus $\mathcal{H}$ has finite variant property [18].

The initial deduction system is given by the tuple $\mathcal{I}_0 = \langle \mathcal{G}, \mathcal{L}_0, \mathcal{H} \rangle$ and we have:

$$\mathcal{L}_0 = \begin{cases} 1: \ x, y \twoheadrightarrow Sig(x, y), \\ 2: \ x, y \twoheadrightarrow Ver(x, y), \\ 3: \ x, y \twoheadrightarrow Bl(x, y), \\ 4: \ x, y \twoheadrightarrow Ubl(x, y). \end{cases}$$

The first step of saturation outputs the following set of deduction rules:

$$\mathcal{L} = \mathcal{L}_0 \cup \begin{cases} 5: \ Sig(x, SK(y)), PK(y) \twoheadrightarrow x, \\ 6: \ Bl(x, y), y \twoheadrightarrow x, \\ 7: \ Sig(Bl(x, y), SK(z)), y \twoheadrightarrow Sig(x, SK(z)). \end{cases}$$

We define a new deduction system $\mathcal{I} = \langle \mathcal{G}, \mathcal{L}, \emptyset \rangle$ and by lemma 1, we have: $t \in \overline{E}^{\mathcal{I}_0}$ iff $t \in \overline{E}^{\mathcal{I}}$ for every set of ground terms E (resp. a ground term t) in normal form. From now we remark that the equational theory employed is the empty one.

Now, let us apply the second step of saturation. The closure applied on rules 1 and 5 outputs the rule 8 : $x, SK(y), PK(y) \twoheadrightarrow x$, the closure applied on rules 3 and 6 outputs the rule 9 : $x, y \twoheadrightarrow x$ which will be deleted by the simplification step introduced above as consequence of lemma 7. The closure applied on rules 1 and 7 outputs the rule 10 : $y, Bl(x, y), SK(z) \twoheadrightarrow Sig(x, SK(z))$.

We prove in the next lemma that the last rule is redundant when the employed equational theory is the empty one.

Lemma 14. *Let* $\mathcal{L}'_1 = \mathcal{L} \cup \{x, SK(y), PK(y) \twoheadrightarrow x\} \cup \{y, BL(x, y), SK(z) \twoheadrightarrow Sig(x, SK(z))\}$ *and let* $\mathcal{L}'_2 = \mathcal{L}'_1 \setminus \{y, BL(x, y), SK(z) \twoheadrightarrow Sig(x, SK(z))\}$. *Suppose that the employed equational theory is the empty one. For any two sets of ground terms in normal form* E *and* F *we have:* $E \twoheadrightarrow^*_{\mathcal{L}'_2} F$ *iff* $E \twoheadrightarrow^*_{\mathcal{L}'_1} F$.

Proof. Let E and F be two sets of normal ground terms. The direct implication is obvious, let us prove the second one. Suppose that $E \twoheadrightarrow^*_{\mathcal{L}'_1} F$ and let us prove that $E \twoheadrightarrow^*_{\mathcal{L}'_2} F$. Suppose that in the $\mathcal{L}'_1$-derivation D starting from E to F there is some steps where the applied rule is in $\mathcal{L}'_1 \setminus \mathcal{L}'_2$ that is, by definition of $\mathcal{L}'_1$ and $\mathcal{L}'_2$, the applied rule is $y, Bl(x, y), Sk(z) \twoheadrightarrow Sig(x, SK(z))$.

Let i be the first step in the derivation where the applied rule is $y, Bl(x, y), Sk(z) \twoheadrightarrow Sig(x, SK(z))$, we prove that this step can be replaced by other steps where the respectives applied rules are in $\mathcal{L}'_2$. $D : E = E_0 \twoheadrightarrow \ldots \twoheadrightarrow E_i \twoheadrightarrow_{y,Bl(x,y),Sk(z)\twoheadrightarrow Sig(x,SK(z))} E_{i+1} \ldots \twoheadrightarrow F$. There is a ground substitution σ in normal form such that $\{y\sigma, Bl(x, y)\sigma, SK(z)\sigma\} \subseteq E_i$ and $E_{i+1} = E_i \cup Sig(x\sigma, SK(z)\sigma)$. Thus, the rule $Bl(x, y), y \twoheadrightarrow x \in \mathcal{L}'_2$ with the substitution σ can be applied first on E_i and outputs $E_{i1} = E_i \cup x\sigma$, then the rule $x, y \twoheadrightarrow Sig(x, y) \in \mathcal{L}'_2$ also with the substitution σ can be applied on E_{i1} and outputs $E_{i1} \cup Sig(x\sigma, SK(z\sigma)) = E_{i+1}$. We deduce that each application of the rule $y, Bl(x, y), Sk(z) \twoheadrightarrow Sig(x, SK(z))$ in D can be replaced by the application of two rules in $\mathcal{L}'_2$. We conclude that $E \twoheadrightarrow^*_{\mathcal{L}'_1} F$ implies $E \twoheadrightarrow^*_{\mathcal{L}'_2} F$. □

Remarks.

Enforcing the termination of the Saturation. The application of the *Saturation* algorithm as is described in section 3 does not terminate. In fact, the rule 10 is an increasing one and *closure* rule can be applied on rules 10 and 7. The application of the closure outputs the rule 11 : $y, y', Bl(Bl(x, y), y'), SK(z) \twoheadrightarrow Sig(x, SK(z))$ which is increasing. We remark that *closure* rule can be applied on the rules 11 and 7 and this application outputs a new increasing rule. In addition, *closure* rule can be applied again on the new obtained rule and the rule 7. We remark also that each such application of closure rule outputs a new increasing rule where the size of the terms in the left hand side is increased and closure rule can be applied again on this new obtained rule and the rule 7. This implies that we have an infinite sequence of application of *closure* rule. We remark that this infinite sequence is due to the presence of the rule 10.

As a consequence from the previous lemma (where we prove that the rule $y, Bl(x,y), SK(z) \twoheadrightarrow Sig(x, SK(z))$ is redundant), we can delete this rule from the system immediately after its creation. This deletion enforces the termination of the *Saturation*.

Saturated deduction system. Let $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$ be the saturated deduction system, we have:

$$\mathcal{L}' = \mathcal{L}_0 \cup \begin{cases} Sig(x, SK(y)), PK(y) \twoheadrightarrow x, \\ Bl(x,y), y \twoheadrightarrow x, \\ Sig(Bl(x,y), SK(z)), y \twoheadrightarrow Sig(x, SK(z)), \\ x, SK(y), PK(y) \twoheadrightarrow x. \end{cases}$$

In $\mathcal{L}'$, we note that only $\mathcal{L}_0$-rules are increasing and the others are decreasing (by definition of *increasing* and *decreasing* rules).

We recall that a derivation D starting from E of goal t is *well-formed* if for all rules $l \twoheadrightarrow r$ applied with substitution σ, for all $u \in l \setminus \mathcal{X}$ we have either $u\sigma \in E$ or $u\sigma$ was constructed by a former decreasing rule.

In the next lemma, we prove that the system $\mathcal{L}'$ satisfies the following lemma.

Lemma 15. *Let E (resp. t) be a set of terms (resp. a term) in normal form such that $t \in \overline{E}^{\mathcal{I}'}$. For all $\mathcal{I}'$-derivations D starting from E of goal t we have either D is well-formed or there is another $\mathcal{I}'$-derivation D' starting from E of goal t such that* $\mathrm{trace}(D) \subseteq \mathrm{trace}(D')$ *and D' is well-formed.*

Proof. We have $t \in \overline{E}^{\mathcal{I}'}$ implies that the set $\Omega(E,t)$ of $\mathcal{I}'$-derivations starting from E of goal t is not empty. Let $D \in \Omega(E,t)$, $D : E = E_0 \twoheadrightarrow E_1 \twoheadrightarrow \ldots \twoheadrightarrow E_{n-1}, t$, we denote $l_i \twoheadrightarrow r_i$ the rule applied at step i with the substitution σ_i: this rule is well-applied if for all $u \in l_i \setminus \mathcal{X}$, we have either $u\sigma \in E$ or $u\sigma$ was obtained by a former decreasing rule, otherwise it is bad-applied.

Suppose that D is not well-formed then there is at least one step in the derivation D where the applied rule is bad-applied. At each such step, one the following rule is applied:

$$\begin{cases} Sig(x, SK(y)), PK(y) \twoheadrightarrow x, \\ Bl(x,y), y \twoheadrightarrow x, \\ Sig(Bl(x,y), SK(z)), y \twoheadrightarrow Sig(x, SK(z)), \end{cases}$$

We note that the rule $x, SK(y), PK(y) \twoheadrightarrow x$ can not be applied at such step because the rules $x \twoheadrightarrow SK(x)$ and $x \twoheadrightarrow PK(x)$ are not in $\mathcal{L}'$.

Let us prove that each application of the first (resp. the second) rule in D such that there is a non variable term in left hand side of the rule where the instance is obtained by a former increasing rule can be deleted from D without altering $\mathrm{trace}(D)$. Let i be the first step where the first (resp. the second) rule is bad applied, that is there is a non variable term in left hand side where the instance is obtained by a former increasing rule. There is only one non variable term in the left hand side of the first (resp. the second) rule which can be obtained by a former increasing rule, this term is $Sig(x, SK(y))$ (resp. $Bl(x,y)$). Since the instance of this term, $Sig(x, SK(y))\sigma$ (resp. $Bl(x,y)\sigma$), is obtained by a former increasing rule this last rule will be $x, y \twoheadrightarrow Sig(x,y)$ (resp.

$x, y \twoheadrightarrow Bl(x, y)$) and let h ($h < i$) be the step where this rule is applied. We deduce that $\{x\sigma, SK(y\sigma)\}$ (*resp.* $\{x\sigma, y\sigma\}$) $\subseteq E_h$ and then the rule applied at step i (which adds $x\sigma$) does not add a new term and the step i can be deleted without modifying in trace(D). Let D' be the obtained derivation, we have trace(D)$'$ = trace(D). We deduce that every step in D' where the rule $Sig(x, SK(y)), PK(y) \twoheadrightarrow x$ (resp. the rule $Bl(x, y), y \twoheadrightarrow x$) is bad applied can be deleted without altering in the trace of D' and let d be the obtained derivation. We note that every application of the rule $Sig(x, SK(y)), PK(y) \twoheadrightarrow x$ (resp. the rule $Bl(x, y), y \twoheadrightarrow x$) in d is a well-application.

Suppose that d is not well-formed then there is at least one step where the rule applied is bad-applied. Let i be the first such step then the rule applied is $Sig(Bl(x, y), SK(z)), y \twoheadrightarrow Sig(x, SK(z))$ and $Sig(Bl(x, y), SK(z))\sigma$ is obtained by a former increasing rule, $x, y \twoheadrightarrow Sig(x, y)$. Let h, ($h < i$), be the step where this increasing rule is applied. We deduce that $\{Bl(x, y)\sigma, SK(z)\sigma\} \subseteq E_h$. If $x\sigma \notin E_i$ then the rule applied at step i in d can be replaced first by the application of $Bl(x, y), y \twoheadrightarrow x$ then the application of $x, y \twoheadrightarrow Sig(x, y)$. Let d' be the obtained derivation, $d' : E \twoheadrightarrow \ldots \twoheadrightarrow E_i \twoheadrightarrow_{Bl(x,y),y\twoheadrightarrow x} E_i, x\sigma \twoheadrightarrow_{x,y\twoheadrightarrow Sig(x,y)} E_i, x\sigma, Sig(x, SK(z))\sigma \twoheadrightarrow \ldots \twoheadrightarrow E_{n-1}, t$. By above and since $x\sigma \notin E_i$ we have either $Bl(x, y)\sigma \in E$ or $Bl(x, y)\sigma$ is obtained by a former decreasing rule.

If $x\sigma \in E_i$ then the rule applied at step i in d can be replaced by the application of $x, y \twoheadrightarrow Sig(x, y)$. Let d'' be the obtained derivation, $d'' : E \twoheadrightarrow \ldots \twoheadrightarrow E_i \twoheadrightarrow_{x,y\twoheadrightarrow Sig(x,y)} E_i, Sig(x, SK(z))\sigma \twoheadrightarrow \ldots \twoheadrightarrow E_{n-1}, t$.

This implies that each bad application of the rule $Sig(Bl(x, y), SK(z)), y \twoheadrightarrow Sig(x, SK(z))$ can be replaced by one (or two) well-applied rules. We deduce that if the derivation D is not well-formed there is another well-formed derivation D'' starting from E of goal t such that trace(D) $\subseteq$ trace(D''). □

We remark that the above lemma is similar to the lemma 3.

In order to solve $\mathcal{I}_0$-ground reachability problems (definition 7), we apply the algorithm defined in section 4. Since the saturation applied on $\mathcal{L}_0$ terminates, by lemmas (4, 5, 8, 9 and 11) we deduce the following corollary:

Corollary 1. *The $\mathcal{I}_0$-ground reachability problem is decidable.*

10 Decidability of reachability problems for subterm convergent theories

In this section, we give a decidability result for the reachability problems for a class of subterm convergent equational theories. We recall that subterm convergent equational theories have finite variant property [18]. The result of this section is entailed by a more general result by Baudet [8], but the proof here in this specific case is much shorter.

We recall that $\mathcal{G}$ is a set of functions symbols and we denote by $\mathcal{H}$ a subterm convergent equational theory and by $\mathcal{I}_0 = \langle \mathcal{G}, \mathcal{L}_0, \mathcal{H} \rangle$ the initial deduction system such that $\mathcal{L}_0$ is the union of functions $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n)$ for some function symbols $f \in \mathcal{G}$.

Definition 9. *(Subterm convergent theories.) An equational theory $\mathcal{H}$ is subterm convergent if it is generated by a convergent rewriting system $\mathcal{R}$ and for each rule $l \to r \in \mathcal{R}$, r is a strict subterm of l.*

In the rest of this section, we give an algorithm to decide the following reachability problem:

$\mathcal{I}_0$-Reachability Problem

Input:	An $\mathcal{I}_0$-constraint system $\mathcal{C}$.
Output:	SAT iff there exists a substitution σ such that $\sigma \models_{\mathcal{I}_0} \mathcal{C}$.

We let $\mathcal{I}' = \langle \mathcal{G}, \mathcal{L}', \emptyset \rangle$ to be the saturated deduction system. We suppose that $r \notin l$ for all rules $l \twoheadrightarrow r \in \mathcal{L}'$ that is rules not satisfying this property will be deleted.

In the following lemma we prove that, in the case of subterm convergent equational theories and under our assumption on the form of initial deduction rules $\mathcal{L}_0$, *Saturation* terminates and the obtained new rules are decreasing.

Lemma 16. *The saturation of $\mathcal{L}_0$ terminates and for every rule $l \twoheadrightarrow r \in \mathcal{L}' \setminus \mathcal{L}_0$ there exists a term $s \in l$ such that r is a strict subterm of s.*

PROOF. Let $l \twoheadrightarrow r \in \mathcal{L}' \setminus \mathcal{L}_0$ and let us prove that this rule satisfies the following property: there is a term $s \in l$ such that $r \in \mathrm{SSub}(s)$. By induction on the number of saturations needed to obtain a rule $l \twoheadrightarrow r$.

Let us first prove this property is true for rules obtained by the step 1 of the saturation. By definition of $\mathcal{H}$, by the fact that variants of term are in normal form and given the assumption that all original rules are $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n)$, this implies:

$$(f(x_1, \ldots, x_n)\theta)\!\downarrow \in \mathrm{SSub}(f(x_1, \ldots, x_n)\theta)$$

Thus, there exists i such that: $(f(x_1, \ldots, x_n)\theta)\!\downarrow \in \mathrm{Sub}(x_i\theta)$

If there is equality, the rule is removed (since $r \notin l$ for all rules $l \twoheadrightarrow r$). This implies that all rules obtained from step 1 of saturation satisfies the property. Since $\mathcal{L}_0$ is finite and since subterm convergent equational theories have finite variant property [18], first step of saturation terminates. Since $u \in \mathrm{SSub}(v)$ implies $u \prec v$, rules obtained by step 1 are decreasing. Let $\mathcal{L}$ be the set of rules obtained by step 1 and let us prove that rules obtained by closure satisfy the property. Let us prove it for the first rule obtained by closure. By definition of closure rule and since rules in $\mathcal{L} \setminus \mathcal{L}_0$ are decreasing, the first closure will be applied on rules $x_1 \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n) \in \mathcal{L}_0$ and $f(s_1, \ldots, s_n), l \twoheadrightarrow r \in \mathcal{L} \setminus \mathcal{L}_0$. Again by definition of closure, the obtained rule is $s_1, \ldots, s_n, l \twoheadrightarrow r$. By definition of decreasing rule, there is a term $u \in \{f(s_1, \ldots, s_n), l\}$ such that $r \in \mathrm{SSub}(u)$, if $u = l$ then the new rule satisfies the property and if $u = f(s_1, \ldots, s_n)$ then there is an integer i such that $r \in \mathrm{Sub}(s_i)$. If $r \in \mathrm{SSub}(s_i)$ the obtained rule satisfies the property else the rule can not be in $\mathcal{L}'$ (since rules $l \twoheadrightarrow r$ with $r \in l$ are deleted). We conclude that the first rule obtained by closure is decreasing and if we apply again closure, it will be applied on a rule in $\mathcal{L}_0$ and a rule not in $\mathcal{L}_0$. We conclude that rules obtained by step 2 satisfy the property and are decreasing . We conclude also that step 2 terminates. □

We recall that increasing rules are of form $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n)$ for a function symbol $f \in \mathcal{G}$ (Lemma 16).

10.1 Decidability result

We recall that our goal is to solve $\mathcal{I}_0$-reachability problem.

Algorithm. Let $\mathcal{C}^0 = ((E_i^0 \rhd v_i^0)_{i\in\{1,\ldots,n\}}, \mathcal{S}^0)$.

> *Step 1.* Guess a finite variant substitution θ for all terms of $\mathcal{C}^0$, apply θ on these terms and normalise them then solve the obtained unification system. Finally, apply the obtained solution α on the constraints. Let $\mathcal{C} = ((E_i \rhd t_i)_{i\in\{1,\ldots,n\}})$ be the obtained constraint system.

We remark that this step terminates and it is also correct (Lemma 5) and complete (Lemma 4). Unless otherwise specified, $\mathcal{I}'$ is the deduction system implicit in all notations in the rest of this section.

We now introduce the notation $\rhd_{inc}$ to denote a deduction constraint that has to be solved using only increasing rules. We say a constraint $E \rhd_{inc} t$ is in solved form if t is a variable. The constraint system is in solved form if all the deduction constraints are in solved form. The application of a decreasing rule $l \twoheadrightarrow r$ on a constraint $E \rhd t$ is defined as follows, and in accordance with Lemma 3:

- let σ be the mgu of the terms in $l \setminus \mathcal{X}$ with a subset F of $E \setminus \mathcal{X}$
- if $\{x_1, \ldots, x_k\} = l \cap \mathcal{X}$, replace $\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta$ with:

$$(\mathcal{C}_\alpha, E \rhd_{inc} x_1, \ldots, E \rhd_{inc} x_k, E \cup \{r\} \rhd t, \mathcal{C}'_\beta)\sigma$$

Where $\mathcal{C}'_\beta$ is constructed from $\mathcal{C}_\beta$ by adding r to each left-hand side. This last construction aims at preserving the inclusion of knowledge sets.

> *Step 2.* Iterate until the constraint system is in solved form or unsolvable:
> 1. Put all tagged deduction constraints $E \rhd_{inc} t$ in solved form;
> 2. If all constraints preceding an untagged $E \rhd t$ are in solved form, Apply non-deterministically $|\mathrm{Sub}(E) \setminus \mathrm{Var}(E)|$ decreasing rules on E. Replace $E \rhd t$ by the obtained deduction constraints, all tagged with *inc*.

Let us prove the completeness and termination of Step 2.

Completeness. The proof of the following lemma is trivial by the form of increasing rules.

Lemma 17. *If $\sigma \models E \rhd_{inc} f(t_1, \ldots, t_n)$ then either $f(t_1, \ldots, t_n)\sigma \in E\sigma$ or $x_1, \ldots, x_n \twoheadrightarrow f(x_1, \ldots, x_n)$ will be in $\mathcal{L}_0$ and for each $i \in \{1, \ldots, n\}$ we have $\sigma \models E \rhd_{inc} t_i$.*

The first part of the iteration consists either in transforming a deduction constraint $E \rhd_{inc} f(t_1, \ldots, t_n)$ into $E \rhd_{inc} t_1, \ldots, E \rhd_{inc} t_n$, or in unifying $f(t_1, \ldots, t_n)$ with $e \in E$. By Lemma 17, given a ground substitution σ such that $\sigma \models E \rhd_{inc} f(t_1, \ldots, t_n)$ there exists a sequence of choices reducing $E \rhd_{inc} f(t_1, \ldots, t_n)$ to a (possibly empty) set of deduction constraints $E\tau \rhd_{inc} u_1, \ldots E\tau \rhd_{inc} u_k$ where the $u_1, \ldots, u_k$ are variables or constants. If there is a constant which is not in $E\tau$ the constraint is not satisfiable (by definition of increasing rules), and the sequence of choices fails.

Let us now consider the second part of the iteration.

Lemma 18. *Assume $\sigma \models E \rhd_{inc} x$ with x the first variable in the sequence of deduction constraints such that $t \in \mathrm{Sub}(x\sigma)$ for some ground term t. Then either there exists $u \in \mathrm{Sub}(E)$ such that $u\sigma = t$ or $t \in \overline{E\sigma}^{\mathcal{L}'_{inc}}$.*

Proof. Let us assume there does not exist $u \in \mathrm{Sub}(E)$ such that $u\sigma = t$. By minimality of x and the determinacy of constraint systems we have $t \notin \mathrm{Sub}(\mathrm{Var}(E)\sigma)$. Since $\mathrm{Sub}(E\sigma) = \mathrm{Sub}(E)\sigma \cup \mathrm{Sub}(\mathrm{Var}(E)\sigma)$ we have $t \notin \mathrm{Sub}(E\sigma)$ and, by hypothesis on x and t, $t \in \mathrm{Sub}(x\sigma)$. Since $\sigma \models E \rhd_{inc} x$ consider a derivation $E_1 = E\sigma \twoheadrightarrow \ldots \twoheadrightarrow E_{n-1} \cup x\sigma$, and let i be minimal such that $t \in \mathrm{Sub}(E_i)$. The index i exists since $t \in \mathrm{Sub}(x\sigma)$, and is different from 1 since $t \notin \mathrm{Sub}(E\sigma)$. By definition of the increasing rules we then must have $E_i = E_{i-1}, t$. □

Consider a $\mathcal{I}'$-constraint system $\mathcal{C} = (\mathcal{C}_\alpha, E \rhd t, \mathcal{C}_\beta)$ satisfied by a substitution σ and all deduction constraints in $\mathcal{C}_\alpha$ are in solved form. By Lemmas 3 and 16 and by the fact that $r \notin l$ for all rules $l \twoheadrightarrow r \in \mathcal{L}'$, all decreasing rules applied on $E\sigma$ yield a term in $\mathrm{Sub}(E\sigma)$. Thus there are at most $|\mathrm{Sub}(E) \setminus \mathrm{Var}(E)|$ different terms that can be obtained by decreasing rule starting from $E\sigma$ and which are not in $\mathrm{Sub}(\mathrm{Var}(E)\sigma)$. Assume a term t is in $\mathrm{Sub}(\mathrm{Var}(E)\sigma) \setminus \mathrm{Sub}(E)\sigma$, and let x be the first variable (in the ordering of deduction constraints) such that $t \in \mathrm{Sub}(x\sigma)$. By definition of constraint systems there exists a deduction constraint $E_x \rhd_{inc} x$ in $\mathcal{C}_\alpha$. Since $E_x \subseteq E$, by Lemma 18, we have $t \in \overline{E_x\sigma}^{\mathcal{L}'_{inc}}$. Again, since $E_x\sigma \subseteq E\sigma$, this implies $t \in \overline{E\sigma}^{\mathcal{L}'_{inc}}$: the decreasing rule was not useful, and can be replaced by a sequence of increasing. Thus in $\overline{E\sigma}$ at most $|\mathrm{Sub}(E) \setminus \mathrm{Var}(E)|$ terms are deducible using decreasing rules. Thus, after a right choice of at most $|\mathrm{Sub}(E) \setminus \mathrm{Var}(E)|$ decreasing rules, all terms deducible from the obtained knowledge set can be deduced using only increasing rules, hence the tagging with *inc* of the final deduction constraint $E \cup \{r_1, \ldots, r_k\} \rhd_{inc} t$, $k = |\mathrm{Sub}(E) \setminus \mathrm{Var}(E)|$.

Termination of Step 2. First let us notice that if a unification is chosen, it unifies two subterms of the constraint system in the empty theory, and thus either the two terms were already equal or it reduces strictly the number of variables in the constraint system. Thus the number of unification choices is bounded by the number of variables in the constraint system. Once all unification have been performed, the termination of the first part of the iteration can

easily be proved by considering the multiset of the right-hand side of the deduction constraints, ordered by the extension to multisets of the (well-founded) subterm ordering. The second part of the iteration obviously terminates. Thus each iteration terminates. Since each iteration decreases strictly the number of non-labelled deduction constraints, Step 2. terminates.

11 Conclusion

In [17], H. Comon-Lundh proposes a two-steps strategy for solving general reachability problems: first, decide ground reachability problems and, second, reduce general reachability problems to ground reachability ones, *e.g.* by providing a bound on the size of a minimal solution of a problem. Our results are in this line: for *contracting* deduction systems, general reachability can be reduced to ground reachability. We strongly conjecture that it permits one to provide a bound on the size of minimal solutions. Also, we leave to the reader the proof of the fact that if saturation terminates, the deduction system is *local* in the sense defined in [9]. Thus, this paper adds a new criterion to the one already known for deciding reachability problems.

In future works, we will investigate how the construction presented here can be extended to equational theories having the finite variant property w.r.t. a non-empty equational theory. We will also try to weaken the definition of $\mu(T)$ for a set of terms T.

References

1. Security Protocols Open Repository. http://www.lsv.ens-cachan.fr/spore/.
2. M. Aadi and A.D. Gordon. A calculus for cryptographic protocols: The spi calculus. *Information and Computation*, pages 148(1):1–70, Jan. 1999.
3. M. Aadi and P. Rogaway. Reconciling two views of cryptography (the computational soundness of formal encryption). *In Proc. 1st IFIP International conference on Theorectical Computer Science (IFIP-TCS), LNCS*, 1872:3–22, Springer–Verlag, 2000.
4. Martín Abadi and Véronique Cortier. Deciding knowledge in security protocols under (many more) equational theories. In *CSFW*, pages 62–76. IEEE Computer Society, 2005.
5. Martín Abadi and Véronique Cortier. Deciding knowledge in security protocols under equational theories. *Theor. Comput. Sci.*, 367(1-2):2–32, 2006.
6. Roberto M. Amadio and Denis Lugiez. On the reachability problem in cryptographic protocols. In Catuscia Palamidessi, editor, *CONCUR*, volume 1877 of *Lecture Notes in Computer Science*, pages 380–394. Springer, 2000.
7. Roberto M. Amadio, Denis Lugiez, and Vincent Vanackère. On the symbolic reduction of processes with cryptographic functions. *Theor. Comput. Sci.*, 290(1):695–740, 2003.
8. Mathieu Baudet. Deciding security of protocols against off-line guessing attacks. In Vijay Atluri, Catherine Meadows, and Ari Juels, editors, *ACM Conference on Computer and Communications Security*, pages 16–25. ACM, 2005.

9. Vincent Bernat and Hubert Comon-Lundh. Normal proofs in intruder theories. In Mitsu Okada and Ichiro Satoh, editors, *ASIAN*, volume 4435 of *Lecture Notes in Computer Science*, pages 151–166. Springer, 2006.
10. Michael Burrows, Martín Abadi, and Roger M. Needham. A logic of authentication. *ACM Trans. Comput. Syst.*, 8(1):18–36, 1990.
11. Yannick Chevalier and Mounira Kourjieh. Key substitution in the symbolic analysis of cryptographic protocols. In Vikraman Arvind and Sanjiva Prasad, editors, *FSTTCS*, volume 4855 of *Lecture Notes in Computer Science*, pages 121–132. Springer, 2007.
12. Yannick Chevalier, Ralf Küsters, Michaël Rusinowitch, and Mathieu Turuani. Deciding the security of protocols with diffie-hellman exponentiation and products in exponents. In Paritosh K. Pandya and Jaikumar Radhakrishnan, editors, *FSTTCS*, volume 2914 of *Lecture Notes in Computer Science*, pages 124–135. Springer, 2003.
13. Yannick Chevalier, Ralf Küsters, Michaël Rusinowitch, and Mathieu Turuani. An np decision procedure for protocol insecurity with xor. *Theor. Comput. Sci.*, 338(1-3):247–274, 2005.
14. Yannick Chevalier, Denis Lugiez, and Michaël Rusinowitch. Towards an automatic analysis of web service security. In Boris Konev and Frank Wolter, editors, *FroCos*, volume 4720 of *Lecture Notes in Computer Science*, pages 133–147. Springer, 2007.
15. Yannick Chevalier and Michaël Rusinowitch. Combining intruder theories. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, *ICALP*, volume 3580 of *Lecture Notes in Computer Science*, pages 639–651. Springer, 2005.
16. J. Clark and J. Jacob. A survey of authentication protocol literature.
17. Hubert Comon-Lundh. Intruder theories (ongoing work). In Igor Walukiewicz, editor, *FoSSaCS*, volume 2987 of *Lecture Notes in Computer Science*, pages 1–4. Springer, 2004.
18. Hubert Comon-Lundh and Stéphanie Delaune. The finite variant property: How to get rid of some algebraic properties. In Jürgen Giesl, editor, *RTA*, volume 3467 of *Lecture Notes in Computer Science*, pages 294–307. Springer, 2005.
19. V. Cortier, S. Delaune, and P. Lafourcade. A survey of algebraic properties used in cryptographic protocols. *Research Report LSV-04-15, LSV, ENS de Cachan*, Sept. 2004.
20. Nachum Dershowitz and Jean-Pierre Jouannaud. Rewrite systems. In *Handbook of Theoretical Computer Science, Volume B: Formal Models and Sematics (B)*, pages 243–320. 1990.
21. Danny Dolev and Andrew Chi-Chih Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–207, 1983.
22. Jean-Marie Hullot. Canonical forms and unification. In Wolfgang Bibel and Robert A. Kowalski, editors, *CADE*, volume 87 of *Lecture Notes in Computer Science*, pages 318–334. Springer, 1980.
23. Steve Kremer and Mark Ryan. Analysis of an electronic voting protocol in the applied pi calculus. In Shmuel Sagiv, editor, *ESOP*, volume 3444 of *LNCS*, pages 186–200. Springer, 2005.
24. G. Lowe. An attack on the needham-schroeder public key authentication protocol. *Information processing letters*, 1995.
25. Gavin Lowe. Casper: A compiler for the analysis of security protocols. In *CSFW*, pages 18–30. IEEE Computer Society, 1997.
26. Gavin Lowe. Towards a completeness result for model checking of security protocols. *Journal of Computer Security*, 7(1), 1999.

27. Alberto Martelli and Ugo Montanari. An efficient unification algorithm. *ACM Trans. Program. Lang. Syst.*, 4(2):258–282, 1982.
28. Catherine Meadows. The nrl protocol analyzer: An overview. *J. Log. Program.*, 26(2):113–131, 1996.
29. Jonathan K. Millen and Vitaly Shmatikov. Constraint solving for bounded-process cryptographic protocol analysis. In *ACM Conference on Computer and Communications Security*, pages 166–175, 2001.
30. J.C. Mitchell, M. Mitchell, and U. Stern. Automated analysis of cryptographic protocols using murϕ. In *Proc. IEEE Symposium on Research in Security and Privacy*, pages 141–153. IEEE Computer Society Press, 1997.
31. Paliath Narendran, Frank Pfenning, and Richard Statman. On the unification problem for cartesian closed categories. *J. Symb. Log.*, 62(2):636–647, 1997.
32. Michaël Rusinowitch and Mathieu Turuani. Protocol insecurity with a finite number of sessions, composed keys is np-complete. *Theor. Comput. Sci.*, 1-3(299):451–475, 2003.