S. Bursuc and H. Comon-Lundh

Protocols, insecurity decision and combination of equational theories

Research Report LSV-02

February, 2009

Laboratoire Spécification et Vérification

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Ecole Normale Supérieure de Cachan
61, avenue du Président Wilson
94235 Cachan Cedex France

Protocols, insecurity decision and combination of equational theories

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

[1] LSV, ENS Cachan & CNRS & INRIA
[2] RCIS, AIST Tokyo

Abstract. We consider the problem of finding attacks for a bounded number of sessions of security protocols. We contribute to this field, showing how to decompose the problem into pieces for a class of equational theories, which includes the hierarchical combinations, as well as non-hierarchical ones. We apply this result to an electronic purse case study: we show the decidability in co-NP of the insecurity problem for a complex equational theory mixing three Abelian groups, exponentiation and homomorphism properties.
The main technical contributions rely on equational logic, term rewriting and combination of theories.

1 Introduction

Security protocols are small distributed programs, which aim at performing some transaction through a unreliable network, while preserving some security property, typically confidentiality. Models of such protocols are transition systems that are both infinitely branching and non-terminating. This yields an undecidable verification problem in general.

Depending on the application area, various approximations or hypotheses may be relevant. Typically, one may assume that messages form a free term algebra (the *perfect cryptography assumption*), yielding a PTIME-decidable transition relation. If we additionally bound the number of times each program is used (then the only source of infinity in the model comes from infinite branching), then the confidentiality problem is co-NP-complete [15]. Perfect cryptography is however a too strong assumption in many applications: sometimes protocols proved secure in that model have simple attacks (see [17] for examples). That is why during the past 6 years several works considered equational properties of primitives (e.g. [5, 9, 13] for some of the first works).

Unfortunately, many protocols rely on complex equational properties of the security primitives, typically arithmetic properties, and some protocols cannot even be executed without considering these properties. This is the case of an electronic purse case study, which we worked on, and whose detailed description is given in [4]. In this example, we need to model some properties of modular exponentiation, including both $(b^x)^y = b^{(x\times y)}$ and $b^x \times b^y = b^{(x+y)}$, and the Abelian group properties of addition and multiplication. Since the security problem encompasses unification modulo the equational theory, we have to be very

careful in the design of the equational theory, as shown in [13]. We use a trick, adding a function symbol for a fixed base exponentiation and restricting the interactions between multiplication and addition, yielding an equational theory, which is sufficient for the protocol execution, yet ruling out a trivial encoding of 10th Hilbert's problem [4]. We show, as an application of the results of the present paper, that the security problem is decidable (in co-NP) for this case study.

We try however to abstract out the particular properties of the equational theory we are considering and give a general result for a class of equational theories. We consider a bounded number of sessions and assume algebraic properties that are described by a convergent rewriting system, modulo the associativity and commutativity of some function symbols. The confidentiality problem is then classically modeled as a *deducibility constraint problem*, which encompasses unification modulo the equational theory. We assume in addition the *finite variant property* [8], which implies the decidability of unification, and which is satisfied by our case study example.

As in previous works, for complex equational theories, we try to break the problem into pieces: our work follows the same lines as [7]. In [7], the authors use a notion of *mode*, which can be automatically computed from the rewrite system and which allows us to reduce the deducibility constraints to some *pure* deducibility constraints, in which the contexts (or proof terms) are well moded. (Though, this is not stated in this way in [7]). For instance, in case of two disjoint theories, there are simply two modes and well-moded proof terms use symbols of one theory only. Hence pure systems can be solved in one single component theory. Now, in the non-disjoint case, the well-moded contexts may use symbols of both theories, but only in a hierarchical way, according to the modes. The problem is that this decomposition might be too rough. For instance, in the extreme case, if we have only one mode, the theory is well-moded, but the hierarchical assumption does not impose any restriction, hence does not help.

It turns out that, in our concrete example, though, after a little massage, we can assign modes to the symbols so that the theory is well-moded, it does not help much since pure systems are still too complex. That is why we need to refine this notion of modes. Roughly, what we do is to assign modes not to function symbols, but rather to contexts.

Our contributions here are twofold: first we give general conditions on equational theories, which allow to reduce general deducibility constraints to pure deducibility constraints and that generalize the well-moded systems. Then, we show a non-trivial instance of this reduction (our case study), solving the pure constraint system in this case.

Our general process of solving deducibility constraints is similar, at least superficially, to combination of decision algorithms [2]: we have to purify the constraints, guess equalities and we take advantage of an ordering on variables. However, our signatures are *not* disjoint and, if we want to state the constraints as a unification problem, we need second order variables.

In section 2 we give our case study rewrite system, recall some definitions and state the finite variant property. We recall in section 3 the definition of deducibility constraints. In section 4 we state a crucial lemma, which corresponds to the "small attack property" in [15]. In section 5, we show how to reduce deducibility constraints to pure ones. In section 6, we solve the pure constraint systems, relying on techniques inspired by the combination of unification algorithms. Many additional proofs are delayed to the appendix, due to lack of space.

2 Terms and rewriting

We use classical notations and terminology from [12] on terms, unification, rewrite systems. For sake of clarity, we recall here some important definitions. $\mathcal{F}$ is a set of function symbol with their arity. $\mathcal{F}_{ac} \subseteq \mathcal{F}$ is a set of associative and commutative symbols. They are considered as varyadic symbols. The sets of terms $\mathcal{T}(\mathcal{F}, X)$ and $\mathcal{T}(\mathcal{F})$ are defined as usual, except that we assumed flattening: these algebras are rather quotients by the associativity and commutativity of some symbols. Everywhere, equality has to be understood modulo AC. Replacements of subterms may require flattening the terms.

A substitution $\sigma = \{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$ maps each variable x_i to t_i and all other variables to themselves. It is extended to an endomorphism of $\mathcal{T}(\mathcal{F}, X)$. We let $\mathrm{Dom}(\sigma) = \{x_1, \ldots, x_n\}$ and $\Im(\sigma) = \{t_1, \ldots, t_n\}$.

Many functions from terms to terms (resp. from terms to sets of terms) are extended without mention to sets of terms by $f(S) \stackrel{\text{def}}{=} \{f(t) \mid t \in S\}$ (resp. $f(S) \stackrel{\text{def}}{=} \bigcup_{t \in S} f(t)$) and to substitutions: $f(\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\} \stackrel{\text{def}}{=} \{x_1 \mapsto f(t_1), \ldots, x_n \mapsto f(t_n)\}$ (resp. $f(\sigma) \stackrel{\text{def}}{=} f(\Im(\sigma))$).

$\mathcal{R}$ is a term rewriting system, which we assume convergent modulo the associativity and commutativity of $\mathcal{F}_{ac}$. $s\!\downarrow$ is the normal form of the term s w.r.t. $\mathcal{R}$ (and modulo AC). In addition, we assume that $\mathcal{R}$ has the *finite variant property*:

Definition 1 ([8]). *A rewrite system $\mathcal{R}$ has the* finite variant property *if, for any term t, there is a (computable) finite set of substitutions $\theta_1, \ldots, \theta_n$ such that, for any substitution σ in normal form, there is an index i and a substitution θ such that $t\sigma\!\downarrow = (t\theta_i)\!\downarrow\theta$ and $\sigma = \theta_i\theta$. $\mathcal{R}$ has the* polynomial variant property *if, in addition, the number of rewriting steps from each $t\theta_i$ to its normal form is bounded by a (fixed) polynomial in the size of t.*

Examples of relevant rewrite systems with the finite (actually polynomial) variant property are given in [8]. We give here another example, which corresponds to our case study whose formalization has been explained in [4].

$\mathcal{F}_{EP} = \{exp, h, +, J_+, e_+, \star, J_\star, e_\star, \bullet, J_\bullet, e_\bullet\}$. For each $\circ \in \{+, \star, \bullet\}$, $\mathcal{R}_{AG(\circ)}$ is the rewrite system modulo AC for $\circ$:

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

In addition, we have the following rewrite rules:

$$\mathcal{R}_0 = \begin{cases} exp(h(x), y) \to h(x \star y) & J_\bullet(h(x)) \to h(J_+(x)) \\ exp(exp(x, y), z) \to exp(x, y \star z) & h(e_+) \to e_\bullet \\ h(x) \bullet h(y) \to h(x + y) & J_\bullet(h(x) \bullet y) \to h(J_+(x)) \bullet J_\bullet(y) \\ h(x) \bullet h(y) \bullet z \to h(x + y) \bullet z & exp(e_\bullet, x) \to h(e_+ \star x) \\ exp(x, e_\star) \to x & \end{cases}$$

Then we define $\mathcal{R}_{EP} = \mathcal{R}_0 \cup \mathcal{R}_{AG(+)} \cup \mathcal{R}_{AG(\star)} \cup \mathcal{R}_{AG(\bullet)}$. These axioms are necessary for the execution of our case study protocol, as explained in [4]. There is no distributivity, otherwise we would get undecidability of our constraint systems.

Lemma 1. *$\mathcal{R}_{EP}$ is convergent (modulo $AC(+, \star, \bullet)$) and has the finite (polynomial) variant property.*

The convergence of $\mathcal{R}_{EP}$ modulo $AC(+, \star, \bullet)$, has been mechanically verified using CiME [11]. The second part of this lemma can be proved using a sufficient condition introduced in [8].

There is a simple procedure, based on narrowing, for the variant computations. This property also implies that unification is decidable and finitary. The polynomial variant property implies that unifiability is in NP:

Corollary 1. *Unification modulo $\mathcal{R}_{EP}$ is finitary and unifiability is in NP.*

$\mathcal{R}_{EP}$ is well-moded, according to [7]. But the best we can do is the following mode assignment: $exp : 1, 0 \to 1$, $h : 0 \to 1$, $\bullet : 1, 1 \to 1$, $\star, + : 0, 0 \to 0$. For this assignment (and hence any other one), the hypotheses of [7] on the large theory (1) are not satisfied; we cannot split the theory.

3 Deducibility Constraint systems

Our constraint systems are identical to former ones (e.g.,[7]). They express the ability for an attacker to get a secret after n interactions with the protocol.

T_1 is the set of terms representing the initial intruder's knowledge and $T_1 \Vdash x_1 \wedge x_1 = s_1$ expresses the ability to deduce from T_1 a message x_1, that will match s_1, the template expected by some agent A. Next, the attacker intercepts the reply t_1 of A, which, together with T_1, yields an increased knowledge T_2, and so on... Variables are put in place of sub-messages that are treated as black-boxes by the participants, and therefore can be replaced with arbitrary messages.

Definition 2 (Constraint systems). *A* deducibility constraint system C *is a finite set of equations S between terms of $\mathcal{T}(\mathcal{F}, X)$, together with a finite sequence of deducibility constraints $\{T_1 \Vdash x_1, \ldots, T_n \Vdash x_n\}$, where each T_i is a finite set of terms in $\mathcal{T}(\mathcal{F}, X)$, $x_1, \ldots, x_n$ are distinct variables and such that:*

Origination: *$Var(T_i) \subseteq \{x_1, \ldots, x_{i-1}\}$, for all $1 \leq i < n$.*
Monotonicity: *$T_i \subseteq T_{i+1}$, for all $1 \leq i < n$.*

Monotonicity expresses that an attacker never forgets the messages. Origination expresses the commitment of the intruder to some message at each protocol step. Note that some variables of S may not belong to $\{x_1, \ldots, x_n\}$.

Example 1. The following is *not* a constraint system, as it violates the origination property:

$$\begin{cases} a, b \Vdash z & z = x + y \\ a, b, x \Vdash y \end{cases}$$

The following is a constraint system (for the EP-rewrite theory):

$$\mathcal{S}_1 = \begin{cases} a \Vdash x & x = h(y + a) \\ a, x \bullet h(b) \Vdash y \end{cases}$$

This corresponds to an agent expecting a message that matches $h(y + a)$ and replying $h(y + a) \bullet h(b)$. Then we ask which values of y can be deduced.

All function symbols in $\mathcal{F}$ are public (i.e. available to the intruder) except some free constants (corresponding to secret data generated by the agents). We let C_{priv} be the set of such constants. $\mathcal{F}$ also contains infinitely many constants out of C_{priv}: this models the ability of the intruder to generate fresh data.

A *recipe* is a term in $\mathcal{T}(\mathcal{F} \setminus C_{priv}, X)$. If T is a finite sequence of terms $t_1, \ldots, t_n$ and ζ is a recipe with variables $x_1, \ldots, x_n$, we often write $\zeta[T]$ for the term $\zeta\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$. Equivalently, recipes can be seen as proof terms corresponding to the deduction of $\zeta[T]\downarrow$ from hypotheses T.

Definition 3 (Solution). *A substitution σ is a* solution of the constraint system C *if its domain contains all the free variables of C and*

- *for every $s = t \in S$, $s\sigma\downarrow = t\sigma\downarrow$*
- *For every $T \Vdash s \in C$, there is a recipe ζ such that $\zeta[T\sigma]\downarrow = s\sigma\downarrow$*

Example 2. Consider the system $\mathcal{S}_1$ of example 1. From $a \Vdash x$, x is assigned a term constructed on a and public function symbols. Then, from $x = h(y + a)$, the same property must hold for y. If $b \in C_{priv}$, we cannot use $x \bullet h(b)$ in the second constraint: the solutions are the substitutions that assign to y any term whose only private constant symbol is a and then assign $h(y\sigma + a)\downarrow$ to x.

We are interested in the satisfiability of constraints systems: "is there an attack on the protocol ?". We cannot expect better than a NP-decision algorithm for this problem. Therefore, in what follows, we will use several NP-reduction steps to supposedly simpler problems.

4 Factors, subterms and conservativity

We prove here a property, called *conservativity*, which roughly states that, if there is a solution σ of C, then there is another solution θ such that $\mathrm{St}(C\theta\downarrow) \subseteq \mathrm{St}(C)\theta\downarrow$, where St is a notion of *semantic* subterm. This corresponds to a "small

attack property" in many previous works on security protocols (starting with [15]): if there is an attack σ, then there is a "small" attack θ which is built only stacking subterms of the protocol rules: its subterms are also (instances of) subterms of the constraint system.

Such a property implies decidability when the deduction system is *local*, there is no AC symbol and the notion of subterm is the usual one. Locality states that all terms occurring in the intruder proofs can be assumed to be subterms of either the hypotheses or the conclusion. It is satisfied by most of (all?) the intruder theories ([10, 7, 4, 1]). Assuming locality and conservativity, we may guess the subterms of C that are deducible, and at which stage, inserting then new constraints. After this transformation, every deducibility constraint must be satisfied by applying a single function symbol, which is easy to decide.

We use the same proof scheme in section 5. The main difference is that we cannot rely on the usual definition of subterm: we would neither get locality nor conservativity. As in [6, 7], we need to consider some blocks of symbols, considered as "pure", as if it was a single symbol. Then the above sketched procedure yields deducibility constraints that must be satisfied by applying a single pure context, instead of a single symbol.

It remains to find the appropriate decomposition into pure blocks of symbols. This is clear when the rewriting system can be split into two rewriting systems whose sets of symbols are disjoint. Similarly, for well-moded systems the modes yield a definition of pure contexts. But we want to go beyond that since, otherwise, we might have to consider any context as pure (e.g., in the EP example), which is not helpful. In order to be as general as possible, we parametrize our results by the notion of (pure) subterms, provided it satisfies a number of hypotheses, which are sufficient for conservativity.

Factors select some (usual, hereafter called "syntactic") subterms in flattened form and the (semantic) set of subterms is defined by $\mathrm{St}(t) = \{t\} \cup \mathrm{St}(\mathrm{Fact}(t))$. *Replacements* are defined, according to St: if $t_1, \ldots, t_k$ are the factors of t and $t = C[t_1, \ldots, t_k]$, then $t[u \mapsto c]$ (the replacement of u with c in t) is defined by: $u[u \mapsto c] \stackrel{\text{def}}{=} c$ and $t[u \mapsto c] \stackrel{\text{def}}{=} C[t_1[u \mapsto c], \ldots, t_k[u \mapsto c]]$ if $t \neq_{(AC)} u$.

Example 3. In case of a combination of two disjoint theories, the function symbols are split in two sets $\mathcal{F}_1$ and $\mathcal{F}_2$ and the factors of a term t whose head symbol is in $\mathcal{F}_1$ are the maximal subterms of t whose head symbols are in $\mathcal{F}_2$.

The second example is our target case study ($top(t)$ is the head symbol of t):

Definition 4 (EP-Factors,$\top$). *We let $Fact_{EP}(t) = Fact_{\top(t)}(t)$ where $Fact_f$ and $\top(t)$ are defined as follows (we let $\circ$ be any symbol in $\{+, \star, \bullet\}$):*

- $\top(t) = \circ$, *if* $top(t) \in \{\circ, e_\circ, J_\circ\}$.
- $\top(h(t)) = \bullet$, *if* $\top(t) = +$ *and* $\top(h(t)) = exp$, *if* $\top(t) = \star$. *In all other cases,* $\top(t) = top(t)$.
- $Fact_\circ(C_\circ[t_1, \ldots, t_n]) = \bigcup_{i=1}^n Fact_\circ(t_i)$ *if* $C_\circ \in \mathcal{T}(\{\circ, J_\circ, e_\circ\}, \mathcal{X})$
- $Fact_\bullet(h(t)) = Fact_+(t)$ *and, in all other cases,* $Fact_\circ(t) = \{t\}$ *if* $\top(t) \neq \circ$

– $Fact_{exp}(exp(u,v)) = Fact_{exp}(u) \cup Fact_\star(v)$, $Fact_{exp}(h(u)) = Fact_\star(u)$ *and, in all other cases,* $Fact_{exp}(t) = \{t\}$ *if* $\top(t) \neq exp$
– *For other function symbols* f, $Fact_f(f(t_1, ..., t_n)) = \{t_1, \ldots, t_n\}$ *and* $Fact_f(t) = \{t\}$ *if* $top(t) \neq f$.

Example 4. – $v = exp(h(a \star b), c \star d)$, $\text{Fact}_{EP}(v) = \{a, b, c, d\}$;
– $v = (a \star b) + h(a \star b + c)$, $\text{Fact}_{EP}(v) = \{h(a \star b + c), a \star b\}$;
– $v = exp(h(a + b), c + d)$, $\text{Fact}_{EP}(v) = \{a + b, c + d\}$;
– $v = h((a + b) \star c) \bullet h(c + d) \bullet (a + d)$, $\text{Fact}_{EP}(v) = \{(a + b) \star c, a + d, c, d\}$.

In addition, we consider two functions F, G from terms to finite sets of terms. F adds some interface contexts: for every u, $F(u) \subseteq \{C_1(u), \ldots, C_n(u)\}$ where $C_1, \ldots, C_n$ are taken out of a fixed set of recipes and G is the inverse of F: $t \in F(u) \Leftrightarrow u \in G(t)$. (Note however that $F(G(t))$ might not be a singleton set.) We assume $\{t\} = F(t) \cap G(t)$. Actually, in case of disjoint or hierarchical combinations, we can choose $F(t) = G(t) = \{t\}$.

Example 5. In the EP example, we let $F(t) = \{t, h(t)\}$, if $top(t) \in \{\star, +, J_\star, J_+\}$; $F(t) = \{t\}$, otherwise. We let G be the inverse of F.

Fact, F, G and the rewrite system $\mathcal{R}$ are assumed to satisfy the following two main properties:

1. For every substitution σ and every term u,

$$\text{Fact}(u\sigma) \subseteq \text{Fact}(u)\sigma \cup \text{Fact}(\sigma) \cup G(\sigma)$$

2. If v is a term whose factors are in normal form, then either $v{\downarrow} \in F(\text{Fact}(v))$ or else $\text{Fact}(v{\downarrow}) \subseteq \text{Fact}(v)$.

In the first condition, the substitution σ can either replace a variable with an alien term (we fall in $\text{Fact}(u)\sigma$), or a term in the same theory as u (we may fall in $\text{Fact}(\sigma)$) or else the term is headed with an interface symbol, and we are in the last case; this is shown by the following

Example 6. In the EP signature, let $u = x \bullet a$ and $x\sigma = h(b \star c)$. Then $\text{Fact}(u) = \{x, a\}$, $\text{Fact}(\sigma) = \{b, c\}$, $\text{Fact}(u\sigma) = \{b \star c, a\}$.

The second property states the main compatibility requirement for the rewriting system. Roughly, either rewriting is compatible with the factoring in theories ($\text{Fact}(v\downarrow) \subseteq \text{Fact}(v)$) or else, rewriting may introduce some interface symbol, and only them, on top of the subterms.

Example 7. Consider again EP. $v = exp(h((a+c) \star J_\star(b)), b)$. $\text{Fact}(v) = \{a+c, b\}$, $v{\downarrow} = h(a + c) \in F(\text{Fact}(v))$ (and $\text{Fact}(v\downarrow) = \{a, c\}$).

We also assume some technical compatibility properties:

3. If v is a term whose factors are in normal form and c is any constant,
   – Either $v[t \mapsto c]{\downarrow} = v{\downarrow}[t \mapsto c]$

- or $v\downarrow = \zeta[t] \in F(t)$ and either $v[t \mapsto c]\downarrow = v\downarrow$ or else $v\downarrow \in F(\text{Fact}(v))$ and $v[t \mapsto c]\downarrow = \zeta[c]$.

4. for any set of terms T, $F(F(T)) = F(T)$ and for any term t, $\text{Fact}(F(t)) = \text{Fact}(G(t)) = \text{Fact}(t)$ and, if t is in normal form, then every $u \in F(T)$ is in normal form.
5. for any constant c, $F(c) = \{c\}$ and if c does not occur in $\mathcal{R}$, then $\forall \zeta.c \in St(\zeta[c])$.
6. If $\text{Fact}(u) = \{u_1, \ldots, u_k\}$, then there exists a recipe ζ such that $\text{Fact}(\zeta) = \{x_1, \ldots, x_k\}$ and $\zeta\{x_1 \mapsto u_1, \ldots, x_k \mapsto u_k\} = u$.

These hypotheses generalize the hierarchical combinations of [7] since, if we take as factors of t the maximal syntactic ill-moded subterms of t, then (appendix C):

Lemma 2. *For well-moded systems, all the hypotheses are satisfied with $F(t) = G(t) = \{t\}$.*

Another relevant example is our target case study (proof in appendix D):

Theorem 1. *EP satisfies our hypotheses.*

With these hypotheses, we get our main lemma:

Lemma 3 (Conservativity). *Let $C = \{T_1 \Vdash v_1, \ldots, T_n \Vdash v_n\} \cup S$ be a constraint system. Then C is satisfiable if and only if it has a solution σ such that $St(C\sigma\downarrow) \subseteq F(G(St(C)\sigma\downarrow))$.*

In words: looking for an attack, it is sufficient to consider substitutions, which are built by stacking pieces of the protocol. This is an analog of the small attack property of, e.g., [15].

The proof of the lemma is tedious and is given in appendix B. The main idea is to replace terms that are in $\text{St}(\sigma) \setminus F(G(\text{St}(C)\sigma\downarrow))$ with an arbitrary constant, yielding again a solution. Roughly, if $\zeta[T]\sigma\downarrow = x\sigma$, then replacing t with c in both members is possible, and yields $\zeta[T]\sigma'\downarrow = x\sigma'$, except when t or some $u \in F(t)$ occurs in the proof, i.e. when, for some $\zeta' \in \text{St}(\zeta)$, $\zeta'[T]\sigma\downarrow \in F(t)$ Then we show that we can construct another proof of the term u, using a recipe in which all private constants, which are irrelevant in the proof $\zeta\sigma\downarrow = u$, are replaced with public constants.

5 Reduction to pure systems

We show in this section that conservativity, locality and the finite variant property, altogether allow to reduce the general problem to pure constraint systems. Locality assumes a function F_1 from terms to finite sets of terms, that adds/removes a fixed set of contexts. It does not need to be the same as F.

Definition 5 (Locality). *A proof system is F_1-*local *(w.r.t. St) if for every term v, recipe ζ and substitution σ in normal form such that $v = \zeta\sigma$, there exist a recipe ζ' such that $\zeta'\sigma\downarrow = v\downarrow$ and for every $u \in St(\zeta')$, $u\sigma\downarrow \in F_1(St(\sigma, v\downarrow))$.*

Definition 6 (Pure systems). *A recipe is* pure *when its factors are variables. A solution of a constraint system C is* pure *if, for every $T \Vdash s \in C$, there is a pure recipe ζ such that $\zeta[T\sigma]\downarrow = s\sigma\downarrow$. A* pure constraint system *is a constraint system whose only pure solutions are considered.*

When subterms are syntactic, pure recipes consist in a single function symbol. For our case study, the pure recipes are listed in lemma 5.

We are now ready for our main reduction theorem:

Theorem 2. *Assume that the notions of factors, subterms and the functions F, G satisfy the hypotheses of the previous section. Assume moreover that the proof system is F_1-local and that the rewrite system has the finite variant property.*

Then the satisfiability of deducibility constraint systems is reducible to the satisfiability of pure constraint systems.

Proof sketch: Let C be the constraint system and σ be a solution of C. By lemma 3, C has a solution σ_1 such that $\mathrm{St}(C\sigma_1\downarrow) \subseteq F(G(\mathrm{St}(C)\sigma_1\downarrow))$.

Using the finite variant property, we get rid of the normalization in this inclusion: there is a variant $C\theta_1\downarrow$ such that $\mathrm{St}(C)\sigma_1\downarrow \subseteq \mathrm{St}(C\theta_1\downarrow)\sigma_2$ for some substitution σ_2 such that $\sigma_1 = \theta_1\sigma_2$. Let $CS_1 = C\theta_1\downarrow$. We have obtained $G(\mathrm{St}(C)\sigma_1\downarrow) \subseteq G(\mathrm{St}(CS_1)\sigma_2)$. We may also take σ_2 out of G by further instanciating the variables of CS_1 using a substitution θ_2 whose range is the fixed finite set of interface contexts: there is some $CS_2 = CS_1\theta_2$ such that $G(\mathrm{St}(CS_1)\sigma_2) \subseteq G(\mathrm{St}(CS_2))\sigma_3$. Now, we let $\overline{F}(u)$ be the "upper bound" of $F(u)$: $\overline{F}(u) = \{C_1(u), \ldots, C_n(u)\}$. All the previous inclusions give us : $\mathrm{St}(C\sigma_1\downarrow) \subseteq F(G(\mathrm{St}(CS_2))\sigma_3) \subseteq \overline{F}(G(\mathrm{St}(CS_2)))\sigma_3$.

By locality, the deducible terms are in $F_1(\mathrm{St}(C\sigma_1\downarrow)) \subseteq F_1(\overline{F}(G(\mathrm{St}(CS_2)))\sigma_3)$. As before, by further instanciating the variables, we may take F_1 out of σ_3, to get $F_1(\mathrm{St}(C\sigma_1\downarrow)) \subseteq \overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3)))\sigma_4$, for some computable θ_3 and $CS_3 = CS_2\theta_3$.

Then, we non-deterministically choose $\theta_1, \theta_2, \theta_3$, add to the equational part of C the equations $x = x\theta_1\theta_2\theta_3$ for each variable of C. Then, we guess which terms in $\overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3))))$ will be deducible by the intruder, and in which order: we insert some deducibility constraints (iteratively) replacing $T_i \Vdash x_i \wedge T_{i+1} \Vdash x_{i+1}$ with a system

$$T_i \Vdash x_i \ \wedge\ T_i \Vdash z_1 \ \wedge\ T_i, z_1 \Vdash z_2 \ \wedge \ldots \wedge\ T_{i+1}, z_1, \ldots, z_m \Vdash x_{i+1}$$

where $z_1, \ldots, z_n$ are new variables and $z_1 = v_1 \wedge \ldots \wedge z_m = v_m$ is added to the equational part, for the guessed $v_1, \ldots, v_m \in \overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3))))$.

By locality, any solution of the original system can be extended to the new variables into a *pure* solution of the resulting system.

Example 8. Consider the deducibility constraint $(a, b, c \in C_{priv})$ in EP:

$$C = \begin{cases} a \star b & \Vdash x \\ a \star b, exp(x, c), J_\star(b \star c), h(a) + c & \Vdash y \end{cases}$$

$\sigma = \{x \mapsto h(a \star b); y \mapsto a\}$ is a solution. Let us see the branch of the above transformation yielding a pure constraint of which an extension of σ is a solution.

First, we compute a variant, guessing that $x = h(z)$ and the normal form of $exp(x, c)$ is $h(z \star c)\theta$. Next, the strict subterms of CS_2 are $a, b, c, z, h(a)$. We guess which of $a, b, c, z, h(a), h(b), h(c), h(z)$ are deducible and in which order. For instance we get the pure system:

$$C' = \begin{cases} a \star b & \Vdash x & x = h(z) \\ a \star b, exp(x, c), J_\star(b \star c), h(a) + c & \Vdash z_1 & z_1 = h(a) \\ a \star b, exp(x, c), J_\star(b \star c), h(a) + c, z_1 & \Vdash z_2 & z_2 = c \\ a \star b, exp(x, c), J_\star(b \star c), h(a) + c, z_1, z_2 & \Vdash y & \end{cases}$$

$\theta = \sigma \uplus \{z_1 \mapsto h(a); z_2 \mapsto c\}$ is a pure solution of C': $z_1\theta$ is obtained using the pure recipe $exp(x_2, x_3)$, $z_2\theta$ is obtained using the pure recipe $x_4 + J_+(x_5)$ and $y\theta$ is obtained using the pure recipe $x_1 \star x_3 \star x_6$. (Each time x_i is replaced with the ith term in T_j).

To conclude our reduction for EP, we are left to prove locality (appendix G):

Lemma 4. *The proof system EP is F'-local with respect to St_{EP}, for $F'(T) = G(T) \cup h(T)$.*

Furthermore, if we assume the polynomial variant property, the above reduction is in NP.

6 Solving pure systems

Now, we are left to solve pure systems. We show in this section that deciding pure systems is in NP in our case study: in the whole section, we assume the signature $\mathcal{F}_{EP}$ and the rewriting system $\mathcal{R}_{EP}$. Though this is a dedicated procedure, the main steps could be reused in other contexts. In summary, it works as follows:

1. Guess which type of recipe is used for each deduction constraint and reduce the possible cases to 3 recipe types only.
2. Guess the top symbols of variables and the equalities between subterms of the constraint system. These two first steps are similar to [7]
3. Reduce further the problem to the case where the top symbols are stable: $\top(u\sigma\downarrow) = \top(u)$ for every solution σ and every subterm u of the constraint. Such a root-stabilization is important in combining decision procedures: it allows to abstract the alien subterms with constants.
4. Eliminate variables from the left hand sides (using the AG properties) and reduce the deduction constraints to linear Diophantine equations
5. Solve a unification problem with additional linear Diophantine equations.

6.1 Reduction to three recipe types

Lemma 5. *In case of EP, any pure and normalized recipe has one of the following forms:*

$$\begin{array}{ll} \zeta_+ : \zeta = \zeta_+[x_1, \ldots, x_n] & \zeta^\star_{exp} : \zeta = exp(y_0, \zeta_\star[x_1, \ldots, x_n]) \\ \zeta_\star : \zeta = \zeta_\star[x_1, \ldots, x_n] & \zeta^+_h : \zeta = h(\zeta_+[x_1, \ldots, x_n]) \\ \zeta^\star_h : \zeta = h(\zeta_\star[x_1, \ldots, x_n]) & \zeta^+_\bullet : \zeta = \zeta_\bullet[x_1, \ldots, x_n] \bullet h(\zeta_+[y_1, \ldots, y_m]), n \geq 1 \end{array}$$

Where $\zeta_\circ$ is a recipe in $\mathcal{T}(\{\circ, J_\circ, e_\circ\}, \mathcal{X})$, for $\circ \in \{+, \star, \bullet\}$.

In the first step of our procedure, we guess, for each constraint $T_i \Vdash x_i \in C$, which of the six above recipe types is used in the deduction. We write $T_i \Vdash_f x_i$, for $f \in \{+, \star, \zeta^\star_{exp}, \zeta^\star_h, \zeta^+_h, \zeta^+_\bullet\}$. Next, we eliminate three cases:

Eliminating $\zeta^\star_{exp}, \zeta^\star_h$ and ζ^+_h.

For each $T_i \Vdash_{\zeta^\star_{exp}} x_i$ do:

1. Guess $u \in T_i$ and add an equation $x_i = exp(u, y_i)$ to the equational part, where y_i is a new variable.
2. Replace $T_i \Vdash_{\zeta^\star_{exp}} x_i$ with $T_i \Vdash_\star y_i$
3. Replace x_i with $exp(u, y_i)$ in every T_j such that $x_i \in var(T_j)$.

We eliminate $\Vdash_{\zeta^+_h}$ and $\Vdash_{\zeta^\star_h}$ in a similar way, by reducing them to $\Vdash_+$ and $\Vdash_\star$, respectively.

Lemma 6. *The above transformations $C \to \bigvee C_j$ are correct and complete.*

In addition, we will split $\zeta^+_\bullet$: the left members of the constraints are now duplicated: $T \Vdash_{\zeta^+_\bullet} x$ becomes $[T; T] \Vdash_\bullet x$ and, by definition, σ is a solution of $[T_1; T_2] \Vdash_\bullet x$ if there are pure recipes $\zeta^1_\bullet$ and ζ^2_+ such that $\zeta^1_\bullet(T_1\sigma) \bullet h(\zeta^2_+(T_2\sigma)) \downarrow = x\sigma$.

After this step, we can assume that C contains only constraints of the type $T_i \Vdash_\circ v_i$, with $\circ \in \{+, \star, \bullet\}$.

6.2 Guessing top symbols and equalities

We enrich the syntax of our constraint system adding constraints $H(u) \in S_t$ (or just $H(u) = f$, when $S_t = \{f\}$). Where S_t is either a finite set of function symbols or "Others".

A (ground, normalized) substitution σ satisfies $H(u) \in S_t$ if $\top(u\sigma\downarrow) \in S_t$, when S_t is a finite set. σ satisfies $H(u) \in$ "*Others*" if $u\sigma$ is a constant, not occurring in $\mathcal{R}_{EP}$. For each variable x in C, we guess a constraint $H(x) = f$. Then we extend $\top$, by letting $\top(x) = f$ if $H(x) = f$.

We guess all equalities between terms in $\mathcal{D}(\mathrm{St}(C)) = \mathrm{St}(C) \cup h(\mathrm{St}(C)) \cup h^{-1}(\mathrm{St}(C))$ (where $h^{-1}(S) = \{t \mid h(t) \in S\}$); guessed equalities are added to the equational part of the constraint. Now, we may only consider solutions σ such that, for any $u, v \in \mathcal{D}(\mathrm{St}(C))$ if $u\sigma\downarrow = v\sigma\downarrow$, then $u = v$ is a consequence of the equational part $\mathcal{S}$ of C: $u =_\mathcal{S} v$.

6.3 Stabilizing the root symbol

If $x, y \in Var(C)$, we let $x \succ_C y$ if, for every constraint $T \Vdash x \in C$, y is a variable of T. This defines an ordering $\succeq_C$, thanks to the origination property. It is extended to symbols of $\mathcal{F}$ by $x \succ_C f$ for $x \in Var(C)$, $f \in \mathcal{F}$ and $f \succ_C h$ for $f \in \mathcal{F} \setminus \{h\}$. Then $\geq$ is the multiset path ordering [12] on the precedence $\succeq_C$.

When the top symbol of a subterm is not stable by substitution and normalization, it equals a smaller subterm of the system:

Lemma 7. *For every $u \in St(C)$ and every solution θ, if $\top(u\theta\downarrow) \neq \top(u)$, there is a $v \in \mathcal{D}(St(C))$, $v < u$ s.t. $u\theta\downarrow = v\theta\downarrow$*

Since equalities have been guessed, we perform the following transformation: for every $u \in \mathrm{St}(C)$ such that $u =_{\mathcal{S}} t$ and $t \in \mathcal{D}(\mathrm{St}(C))$ and $u > t$, then replace u with t in the left sides of deducibility constraints.

By the lemma 7, after iterating the above replacements, the root symbols are stable, and the resulting system is still a constraint system.

Lemma 8. *The above transformation $C \to C'$ preserves the solutions and, for every solution σ of C', for every $u \in St(C')$, $\top(u\sigma\downarrow) = \top(u)$.*

Example 9.

$$\left.\begin{array}{ll} a \bullet b, c & \Vdash_+ x \\ a \bullet b, c, x - c & \Vdash_\bullet y \end{array}\right\} \Rightarrow \left\{\begin{array}{l} a \bullet b, c \Vdash_+ x \\ a \bullet b, c \Vdash_\bullet y \end{array}\right.$$

Assume $x\sigma = a \bullet b + c$. Then the equality $x - c = a \bullet b$ is part of the equations of C and, since $x - c > a \bullet b$, the transformation yields the system on the right.

6.4 Eliminating variables from left hand sides: reducing deducibility constraints to linear Diophantine equations

At this stage, we could consider the alien subterms in the T_i's as constants and translate the deducibility constraints into a Diophantine systems. This yields however non-linear Diophantine equations, as shown in [3]: we need to rely on monotonicity and the Abelian group properties for further simplifications.

Let $T_i \Vdash_\circ x_i$ be a constraint of C. In this step, we eliminate variables from $Fact_\circ(T_i)$, if they are headed with $\circ$ and, at the same time, introduced by a $\circ$ constraint. We will show in lemma 9 that the other case, when $H(x) = \circ$ and $x\sigma$ is constructed by a different rule $\circ' \neq \circ$, is already handled in step 6.3.

First let $\circ \in \{+, \star\}$. For $i = 1$ to n (i.e. in increasing order on left sides of deducibility constraints) do:

If $T_i \Vdash_\circ x_i$ then let $\mathcal{X}_i^\circ$ be the set of variables x_j, $j < i$ such that $H(x_j) = \circ \in C$ and $T_j \Vdash_\circ x_j \in C$, in:

For every $u \in T_i$ let $u = \zeta_\circ[x_{i_1}, \ldots, x_{i_k}, u_1, \ldots, u_m]$ be such that $\mathrm{Fact}_\circ(u) \cap \mathcal{X}_i^\circ = \{x_{i_1}, \ldots, x_{i_k}\}$

Replace u with $\overline{u}^\circ = \zeta_\circ[e_\circ, \ldots, e_\circ, u_1, \ldots, u_m]\downarrow$ in T_i.

Example 10.

$$\left.\begin{array}{ll} a & \Vdash_+ x \\ a, b + x & \Vdash_+ y \end{array}\right\} \implies \left\{\begin{array}{ll} a & \Vdash_+ x \\ a, b & \Vdash_+ y \end{array}\right.$$

This preserves the solutions since, from any recipe ζ_+ such that $\zeta_+(a, b+x\sigma) \downarrow = y\sigma$, we can build a new recipe ζ'_+ such that $\zeta'_+(a, b) = y\sigma$, thanks to the monotonicity of the constraints (ζ'_+ may rebuild $x\sigma$) and the Abelian group properties of $+$ (we can build ζ_+ from ζ'_+ by subtracting $x\sigma$ when necessary). These ideas were already applied to deduction constraints modulo Abelian groups in [16].

Now let $\circ = \bullet$ and $[T_i, T_i] \Vdash_\bullet x_i \in C$. For the second copy of T_i, we apply the same procedure as above, using $\circ = +$. For the first copy (used by a $\zeta_\bullet$ recipe), we do the following: for each $u \in T_i$, let $u = u' \bullet h(u'')$ (with u' or u'' possibly empty). Then $\overline{u} = \overline{u'}^\bullet \bullet h(\overline{u''}^+)$, where $\overline{u'}^\bullet$ and $\overline{u''}^+$ are obtained from u', respectively u'', by applying the above procedure with $\circ = \bullet$, respectively $\circ = +$. For instance, $x \bullet a \bullet h(\overline{y + b}) = a \bullet h(b)$, if $H(x) = \bullet$ and $H(y) = +$.

Note that we use and may loose monotonicity here, but we no longer need it in further sections.

Lemma 9. *The above transformations are correct and complete and in the resulting constraint system C':*

- *if $T_i \Vdash_\circ x_i \in C'$ and $\circ \in \{+, \star\}$, then for every $x \in Fact_\circ(T_i)$, $H(x) \neq \circ$.*
- *if $[T'_i, T''_i] \Vdash_\bullet x_i \in C'$, then:*
  - *for every $x \in Fact_+(T''_i)$, $H(x) \neq +$.*
  - *for every $u = u_1 \bullet \ldots u_n \bullet h(v_1 + \ldots + v_m) \in T'_i$ such that $Fact_\bullet(u) = \{u_1, \ldots, u_n, v_1, \ldots, v_m\}$, for every $x \in \mathcal{X}$, $x \in \{u_1, \ldots, u_n\} \Rightarrow H(x) \neq \bullet$ and $x \in \{v_1, \ldots, v_m\} \Rightarrow H(x) \neq +$.*

6.5 Turning deduction constraints into linear Diophantine equations

We do the following transformation of deducibility constraints into equations. This transformation is possible and correct by lemmas 7 and 9.

$$\sum_i t_i^1, \ldots, \sum_i t_i^n \Vdash_+ x \quad \implies \quad x = \sum_{i,j} \lambda_j t_i^j$$

if, $\forall i, j. \top(t_i^j) \neq +$ and $\lambda_1, \ldots, \lambda_n$ are new formal integer variables, representing the number of times each term is selected in the recipe. There is a similar rule for $\Vdash_\star$.

For $\Vdash_\bullet$, we use here a multiplicative notation for $\bullet$ and an additive one for $+$:

$$[\ (\textstyle\prod_i t_i^1) \bullet h(\sum_i u_i^1),\ \ldots,\ (\prod_i t_i^n) \bullet h(\sum_i u_i^n)\ ;\ \sum_i v_i^1, \ldots, \sum_i v_i^m\ ] \ \Vdash\ x$$
$$\Rightarrow \begin{cases} x = x_1 \bullet h(x_2) \\ x_1 = \prod_{j=1}^n \prod_i (t_i^j)^{\lambda_j} \\ x_2 = \sum_{j=1}^n \sum_i \lambda_j u_i^j + \sum_{j=1}^m \sum_i \mu_j v_i^j \end{cases}$$

If $\forall i, j. \top(t_i^j) \notin \{\bullet, h\}$ & $\top(u_i^j), \top(v_i^j) \neq +$ and $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m$ are integer variables.

Example 11.

$$[\ a \bullet h(2b),\ a^3 \bullet b\ ;\ 2a+3b, 2b\ ] \Vdash\ x$$

is turned into $x = x_1 \bullet h(x_2)$, $x_1 = a^{\lambda_1} \bullet a^{3\lambda_2} \bullet b^{\lambda_2} = a^{\lambda_1+3\lambda_2} \bullet b^{\lambda_2}$, $x_2 = 2\lambda_1 b + 2\mu_1 a + 3\mu_1 b + 2\mu_2 b = 2\mu_1 a + (2\lambda_1 + 3\mu_1 + 2\mu_2)b$.

6.6 Solving the system of equations

We have now to solve a system of equations $E = E_1 \cup E_2$, where E_1 contains equations of the form $x = \Sigma_\circ \lambda_i \alpha_i t_i, H(t_i) \neq \circ$ and E_2 is a set of usual equations. After applying the finite variant property, our procedure for solving E is similar to unification procedures in the union of disjoint theories, except that we have in addition the linear Diophantine equations coming from the deducibility constraints. We recall here very briefly the main steps.

Step 1: apply the finite variant property. Equations modulo EP are reduced to equations in a combination of 3 AC theories.

Step 2: guess equalities, theories and an occurrence ordering. Since new equalities can be introduced in step 1, we guess once more the equalities between the subterms of E.

Step 3: turn the system into linear diophantine equations. The equations modulo AC yield linear Diophantine systems. That is where constraints of E_1 are inserted. After the above steps, equations in E_1 have the following form:

$$\beta_1 u_1 \circ \ldots \circ \beta_k u_k = \lambda_1 \alpha_1 t'_1 \circ \ldots \circ \lambda_n \alpha_n t'_n,\ \forall i. H(t'_i) \neq \circ$$

Since equalities have been guessed, we can simplify the equations and turn it into a linear system. For instance, if $\circ = +$:

$$\alpha_1 \lambda_1 = \beta_1 \mu_1^1 + \ldots + \beta_m \mu_1^m + \gamma_1\ \wedge\ \ldots\ \wedge\ \alpha_n \lambda_n = \beta_1 \mu_n^1 + \ldots + \beta_m \mu_n^m + \gamma_n$$

From there, we simply go on with the combination of unification procedures.

7 Conclusion

Gathering together all previous steps, we get an NP decision procedure for pure constraints systems, hence, thanks to theorems 1 and 2, we get

Theorem 3. *The insecurity problem for EP is in NP.*

This achieves our main goal: we get a decision procedure for our case study, relying on a general combination technique for deducibility constraints in a non-disjoint case.

It remains however (this is future work) to show how to compute from a rewriting system the functions $Fact$, F, G, which satisfy the required hypotheses. In the present paper, we have only shown how to do it in our case study (or in well-moded systems). This is a very ambitious program: from an AC-convergent rewrite system, find automatically a splitting, which is amenable to combination results.

References

1. S. Anantharaman, P. Narendran, and M. Rusinowitch. Intruders with caps. In *Proc. 18th Int. Conf. on Rewriting Techniques and Applications (RTA)*, volume 4533 of *Lecture Notes in Computer Science*, 2007.
2. F. Baader and K. U. Schulz. Unification in the union of disjoint equational theories: Combining decision procedures. *J. Symb. Comput.*, 21(2):211–243, 1996.
3. S. Bursuc, H. Comon-Lundh, and S. Delaune. Associative-commutative deducibility constraints. In *Proc. 24th Symp. on Theoretical Aspects of Computer Science (STACS'07)*, volume 4393 of *Lecture Notes in Computer Science*, pages 634–645, 2007.
4. S. Bursuc, H. Comon-Lundh, and S. Delaune. Deducibility constraints, equational theory and electronic money. In *Rewriting, Computation and Proof — Essays Dedicated to Jean-Pierre Jouannaud on the Occasion of his 60th Birthday*, volume 4600 of *Lecture Notes in Computer Science*. Springer, 2007.
5. Y. Chevalier, R. Kuester, M. Rusinowitch, and M. Turuani. An NP decision procedure for protocol insecurity with xor. In Kolaitis [14].
6. Y. Chevalier and M. Rusinowitch. Combining Intruder Theories. In *Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings*, volume 3580 of *Lecture Notes in Computer Science*, pages 639–651. Springer, 2005.
7. Y. Chevalier and M. Rusinowitch. Hierarchical combination of intruder theories. In *Proc. Rewriting Techniques and Application*, volume 4098 of *Lecture Notes in Computer Science*, pages 108–122, 2006.
8. H. Comon-Lundh and S. Delaune. The finite variant property: How to get rid of some algebraic properties. In *Proc. of the 16th International Conference on Rewriting Techniques and Applications (RTA'05)*, volume 3467 of *Lecture Notes in Computer Science*, 2005.
9. H. Comon-Lundh and V. Shmatikov. Intruder deductions, constraint solving and insecurity decision in preence of exclusive or. In Kolaitis [14].
10. H. Comon-Lundh and R. Treinen. Easy intruder deductions. In *Verification: Theory and Practice, Essays Dedicated to Zohar Manna on the Occasion of His 64th Birthday*, volume 2772 of *Lecture Notes in Computer Science*, pages 225–242. Springer-Verlag, 2003.
11. E. Contejean and C. Marché. CiME: Completion modulo E. In *Proc. Rewriting Techniques and Applications*, volume 1103 of *Lecture Notes in Computer Science*, pages 416–419, 1996.
12. N. Dershowitz and J.-P. Jouannaud. Rewrite systems. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 243–309. 1990.
13. D. Kapur, P. Narendran, and L. Wang. An e-unification algorithm for analyzing protocols that use modular exponentiation. In *Proc. Rewriting Techniques and Applications*, volume 2706 of *Lecture Notes in Computer Science*, 2003.
14. P. Kolaitis, editor. *Eighteenth Annual IEEE Symposium on Logic in Computer Science*, Ottawa, Canada, June 2003. IEEE Computer Society.
15. M. Rusinowitch and M. Turuani. Protocol insecurity with finite number of sessions is np-complete. In *Proc.14th IEEE Computer Security Foundations Workshop*, Cape Breton, Nova Scotia, June 2001.
16. V. Shmatikov. Decidable analysis of cryptographic protocols with products and modular exponentiation. In *Proc. European Symposium on Programming (ESOP'04)*, volume 2986 of *Lecture Notes in Computer Science*. Springer-Verlag, 2004.

17. Security protocols open repository. http://www.lsv.ens-cachan.fr/spore/.

A Polynomial variant property and unifiability

Corollary 1 *Unification modulo $\mathcal{R}_{EP}$ is finitary and unifiability is in NP.*

Proof. Let P be a unification problem. The finite variant property ensures that there is only a finite number of variants P_i (whose size is polynomial w.r.t. the size of P) of P and for every substitution σ, there is an index i and a substitution θ such that $P\sigma\downarrow = P_i\theta$. We can guess the variant P_i that is satisfiable and verify this fact in polynomial time (unifiability for $AC(+,*,\bullet)$ is in NP). Thus, we have that unifiability for $\mathcal{R}_{EP}$ is also in NP. Moreover, since unification for $AC(+,*,\bullet)$ is finitary, we deduce that this is also the case for the theory $\mathcal{R}_{EP}$.

B Proof of conservativity (Lemma 3)

Before giving the proof, let us give some examples:

Example 12 (Conservativity).

1. Let $\mathcal{C}$ be
$$\begin{cases} h(a+b)\bullet c & \Vdash x \\ h(a+b)\bullet c, exp(x\bullet J_\bullet(c), c\star d) & \Vdash y \\ y = h(z\star d) & \end{cases}$$
The minimal solution of this system is $\sigma = \{x \mapsto h(a+b)\bullet c, y \mapsto h((a+b)\star c\star d), z \mapsto (a+b)\star c\}$. The term $a+b \in \mathrm{St}(C\sigma\downarrow)$ is in $G(\mathrm{St}(C)\sigma\downarrow)\backslash\mathrm{St}(C)\sigma\downarrow$. That is why we need G for conservativity.
2. Let $\mathcal{C}$ be
$$\begin{cases} h((a+b)\star c) & \Vdash x \\ h((a+b)\star c), exp(x, J_\star(c))\star c\star d & \Vdash y \\ y = z\star d & \end{cases}$$
The minimal solution of this system is $\sigma = \{x \mapsto h((a+b)\star c), y \mapsto h(a+b)\star c\star d, z \mapsto h(a+b)\star c\}$. The term $h(a+b) \in \mathrm{St}(C\sigma\downarrow)$ is in $F(\mathrm{St}(C)\sigma\downarrow)\backslash\mathrm{St}(C)\sigma\downarrow$. That is why we need F for conservativity.
3. Let $\mathcal{C}$ be
$$\begin{cases} h(a\star b\star c) & \Vdash x \\ h(a\star b\star c), exp(x, J_\star(c))\bullet c\bullet d & \Vdash y \\ y = z\bullet d & \end{cases}$$
The minimal solution of this system is $\sigma = \{x \mapsto h(a\star b\star c), y \mapsto h(a\star b)\bullet c\bullet d, z \mapsto h(a\star b)\bullet c\}$. The term $a\star b \in \mathrm{St}(C\sigma\downarrow)$ is in $G(\mathrm{St}(C)\sigma\downarrow)\backslash\mathrm{St}(C)\sigma\downarrow$. That is why we need G for conservativity.

4. Let $\mathcal{C}$ be
$$\begin{cases} h(a \star b + c) & \Vdash x \\ h(a \star b + c), x \bullet h(J_+(c)) \star c \star d & \Vdash y \\ y = z \star d \end{cases}$$
The minimal solution of this system is $\sigma = \{x \mapsto h(a \star b + c), y \mapsto h(a \star b) \star c \star d, z \mapsto h(a \star b) \star c\}$. The term $h(a \star b) \in \mathrm{St}(C\sigma\downarrow)$ is in $F(\mathrm{St}(C)\sigma\downarrow) \backslash \mathrm{St}(C)\sigma\downarrow$. That is why we need F for conservativity.

We start with a series of preliminar technical lemmas.

Lemma 10. *If ζ is a recipe with $Fact(\zeta) = \{x_1, \ldots, x_k\} \subseteq X$ and $t, v_1, \ldots, v_k$ are terms s.t. $\zeta\{x_1 \mapsto v_1, \ldots, x_k \mapsto v_k\} \neq t$, then $\zeta\{x_1 \mapsto v_1, \ldots, x_k \mapsto v_k\}[t \mapsto c] = \zeta[x_1 \mapsto v'_1, \ldots, x_1 \mapsto v'_k]$ where*

- *$v'_i = v_i[t \mapsto c]$ if $v_i \notin F(t)$*
- *$v'_i \in \{v_i, \chi[c]\}$ otherwise. Where χ is s.t. $v_i = \chi[t] \in F(t)$.*

Proof. From hypothesis 1,

$$\mathrm{Fact}(\zeta\{x_1 \mapsto v_1, \ldots, x_k \mapsto v_k\} \subseteq \{v_1, \ldots, v_k\} \cup \mathrm{Fact}(v_1, \ldots, v_k) \cup G(v_1, \ldots, v_k)$$

Moreover, $t \in \mathrm{St}(G(v_i))$ iff $t \in \mathrm{St}(v_i)$ or $v_i \in F(t)$ since, by hypothesis 4, for any v, $\mathrm{Fact}(F(v)) = \mathrm{Fact}(v)$ hence for any u, $\mathrm{Fact}(G(u)) = \mathrm{Fact}(u)$. It follows that, for every i, either $v_i \notin F(t)$, in which case $\zeta\{x_i \mapsto v_i\}[t \mapsto c] = \zeta\{x_i \mapsto v_i[t \mapsto c]\}$, or else $v_i \in F(t)$. In the latter case, either $t \notin \mathrm{St}(\zeta\{x_i \mapsto v_i\})$, in which case $\zeta\{x_i \mapsto v_i\}[t \mapsto c] = \zeta\{x_i \mapsto v_i\}$, or else $v_i = \chi[t]$ and $\zeta\{x_i \mapsto v_i\}[t \mapsto c] = \zeta\{x_i \mapsto v'_i\}$, where $v'_i = \chi[c]$.

Lemma 11. *Let u, t be terms in normal form and σ be a normalized substitution s.t. $t \notin G(St(u)\sigma\downarrow)$. Let c be a constant not occurring in $t, u, \sigma, \mathcal{R}$. Then we have:*

1. *$(u\sigma)[t \mapsto c] = u(\sigma[t \mapsto c])$ and $t \in St(u\sigma) \implies t \in St(\sigma)$.*
2. *$(u\sigma)[t \mapsto c]\downarrow = u\sigma\downarrow[t \mapsto c] = u(\sigma[t \mapsto c])\downarrow$.*
3. *$t \in St(u\sigma\downarrow) \implies t \in St(\sigma)$.*

Proof. **1.** We proceed by induction on the size of u. If u is a variable, the result is immediate.

Otherwise, let $\zeta[x_1, \ldots, x_k], u_1, \ldots, u_k$ be s.t. $u = \zeta[u_1, \ldots, u_k]$ and $Fact(u) = \{u_1, \ldots, u_k\}$. From hypothesis 1, we have:

$$Fact(\zeta[u_1\sigma, \ldots, u_k\sigma]) \subseteq \{u_1\sigma, \ldots, u_k\sigma\} \cup Fact(u_1\sigma, \ldots, u_k\sigma) \cup G(u_1\sigma, \ldots, u_k\sigma)$$

Moreover, by hypothesis 4, $t \in St(G(u_i\sigma))$ iff $t \in St(u_i\sigma)$ or $u_i\sigma \in F(t)$. Therefore, $t \in St(u\sigma)$ iff $t \in \{u\sigma\} \cup St(u_1\sigma, \ldots, u_k\sigma) \cup G(u_1\sigma, \ldots, u_k\sigma)$. By hypothesis of the lemma, for every i, $u_i\sigma\downarrow \notin F(t)$ and $u\sigma\downarrow \notin F(t)$. Then, by hypothesis 4, $u\sigma \notin F(t)$ and, for every i, $u_i\sigma \notin F(t)$. It follows that $t \in St(u\sigma)$ iff $t \in St(u_1\sigma, \ldots, u_k\sigma)$. Moreover, since, for all i, $u_i\sigma \notin F(t)$, by lemma 10, we get:

$$(u\sigma)[t \mapsto c] = \zeta[(u_1\sigma)[t \mapsto c], \ldots, (u_k\sigma)[t \mapsto c]]$$

By induction hypothesis, for every i, $(u_i\sigma)[t \mapsto c] = u_i(\sigma[t \mapsto c])$ and $t \in \mathrm{St}(u_i\sigma) \implies t \in St(\sigma)$. Then we conclude $(u\sigma)[t \mapsto c] = u(\sigma[t \mapsto c])$ and $t \in \mathrm{St}(u\sigma) \implies t \in \mathrm{St}(\sigma)$.

2. Let us prove the first equality. We proceed by induction on the derivation of $u\sigma$ to its normal form. If $u\sigma$ is in normal form, the result follows immediately.

Suppose now that there is at least one rewriting step. Let $Fact(u) = \{u_1, \ldots, u_k\}$ and $u = \zeta[u_1, \ldots, u_k]$. Consider the bottom-up normalization of $u\sigma$:

$$u\sigma\downarrow = \zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow]\downarrow$$

Note that, for each i, u_i satisfies the conditions of the lemma and we can apply the induction hypothesis for it: $(u_i\sigma)[t \mapsto c]\downarrow = u_i\sigma\downarrow[t \mapsto c]$.

As before, $u_i\sigma \notin F(t)$ and $u_i\sigma\downarrow \notin F(t)$, since $t \notin G(St(u)\sigma\downarrow)$ and any term in $F(t)$ is in normal form (by hypothesis 4). It follows that $\mathrm{Fact}(\zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow]) \subseteq \mathrm{St}(u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow)$. Since, in addition, $\zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow] \neq t$ (since t is in normal form and $(u\sigma)\downarrow \neq t$), we get

$$\zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow][t \mapsto c] = \zeta[u_1\sigma\downarrow[t \mapsto c], \ldots, u_k\sigma\downarrow[t \mapsto c]]$$

Moreover, as shown above, $\mathrm{Fact}(\zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow]) \subseteq \mathrm{St}(u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow)$: all factors of $\zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow]$ are in normal form and $u\sigma\downarrow \notin F(t)$. Then, by hypothesis 3,

$$u\sigma\downarrow[t \mapsto c] = \zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow][t \mapsto c]\downarrow$$

Gathering together the identities:

$$u\sigma\downarrow[t \mapsto c] = \zeta[u_1\sigma\downarrow[t \mapsto c], \ldots, u_k\sigma\downarrow[t \mapsto c]]\downarrow$$

By induction hypothesis, $(u_i\sigma)[t \mapsto c]\downarrow = (u_i\sigma)\downarrow[t \mapsto c]$. Moreover, we have seen in point 1. above that $\zeta[u_1\sigma, \ldots, u_k\sigma][t \mapsto c] = \zeta[u_1\sigma[t \mapsto c], \ldots, u_k\sigma[t \mapsto c]]$. Then

$$\begin{aligned}
(u\sigma)\downarrow[t \mapsto c] &= \zeta[u_1\sigma\downarrow[t \mapsto c], \ldots, u_k\sigma\downarrow[t \mapsto c]]\downarrow \\
&= \zeta[(u_1\sigma)[t \mapsto c]\downarrow, \ldots, (u_k\sigma)[t \mapsto c]\downarrow]\downarrow \\
&= \zeta[(u_1\sigma)[t \mapsto c], \ldots, (u_k\sigma)[t \mapsto c]]\downarrow \\
&= \zeta[u_1\sigma, \ldots, u_k\sigma][t \mapsto c]\downarrow \\
&= (u\sigma)[t \mapsto c]\downarrow
\end{aligned}$$

The second equality follows from $(u\sigma)[t \mapsto c] = u(\sigma[t \mapsto c])$, the first point of the lemma.

3. We proceed by induction on the derivation of $u\sigma$ to its normal form. If $u\sigma$ is in normal form, the result follows by the first part of the lemma.

Suppose now that there is at least one rewriting step. Let $Fact(u) = \{u_1, \ldots, u_k\}$ and $u = \zeta[u_1, \ldots, u_k]$. Consider the bottom-up normalization of $u\sigma$:

$$u\sigma\downarrow = \zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow]\downarrow$$

As before $\mathrm{Fact}(\zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow]) \subseteq \mathrm{St}(u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow)$. Then

$$\begin{aligned}
\mathrm{St}(u\sigma\downarrow) &= \{u\sigma\downarrow\} \cup \mathrm{St}(\mathrm{Fact}(u\sigma\downarrow)) \\
&\subseteq \{u\sigma\downarrow\} \cup \mathrm{St}(\mathrm{Fact}(\zeta[u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow])) \text{ by hyp. 2} \\
&\subseteq \{u\sigma\downarrow\} \cup \mathrm{St}(u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow)
\end{aligned}$$

So, if $t \in \mathrm{St}(u\sigma\downarrow)$, since $t \neq u\sigma\downarrow$, then $t \in \mathrm{St}(u_1\sigma\downarrow, \ldots, u_k\sigma\downarrow)$, hence $t \in \mathrm{St}(\sigma)$ by induction hypothesis.

Lemma 12. *Let v be a term with all its factors in normal form and u, t be terms in normal form. Then we have:*

- *if $v\downarrow \notin F(t)$, then $v[t \mapsto u]\downarrow = v\downarrow[t \mapsto u]\downarrow$*
- *if $v\downarrow \in F(t)$, then either $v[t \mapsto u]\downarrow = v\downarrow$ or else $v[t \mapsto u]\downarrow = \zeta[u]\downarrow$, for some ζ s.t. $v\downarrow = \zeta[t] \in F(t)$.*

Proof. Let c be a constant not-occurring in $v, t, \mathcal{R}$. $v[t \mapsto u] = v[t \mapsto c][c \mapsto u]$ since all occurrences of c in $v[t \mapsto c]$ are in $\mathrm{St}(v[t \mapsto c])$, by hypothesis 5. By convergence of $\mathcal{R}$,

$$v[t \mapsto u]\downarrow = v[t \mapsto c]\downarrow[c \mapsto u]\downarrow.$$

By hypothesis 3, either $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$ or else $v\downarrow \in F(t)$. In the first case, $v[t \mapsto u]\downarrow = v\downarrow[t \mapsto c][c \mapsto u]\downarrow = v\downarrow[t \mapsto u]\downarrow$. In the second case, again by hypothesis 3, either $v[t \mapsto c]\downarrow = v\downarrow$ or else $v\downarrow \in F(\mathrm{Fact}(v))$, $v\downarrow = \zeta[t]$ and $v[t \mapsto c]\downarrow = \zeta[c]$. In the first case, $v[t \mapsto u]\downarrow = v\downarrow[c \mapsto u]\downarrow = v\downarrow$ (since c does not occur in $v\downarrow$). In the second case, let $v\downarrow = \zeta[t]$, then, by hypothesis 5,

$$v[t \mapsto u]\downarrow = \zeta[c][c \mapsto u]\downarrow = \zeta[u]\downarrow$$

Putting lemmas 11 and 12 together we get the following:

Corollary 2. *Let v be a term, t, u be terms in normal form and σ be a normalized substitution. Assume that either all factors of $v\sigma$ are in normal form or $t \notin G(St(v)\sigma\downarrow)$. Then:*

- *if $v\sigma\downarrow \notin F(t)$, then $(v\sigma)[t \mapsto u]\downarrow = v\sigma\downarrow[t \mapsto u]\downarrow$.*
- *if $v\sigma\downarrow \in F(t)$, then either $(v\sigma)[t \mapsto u]\downarrow = v\sigma\downarrow$ or else $(v\sigma)[t \mapsto u]\downarrow = \zeta[u]\downarrow$ for some ζ s.t. $v\sigma\downarrow = \zeta[t] \in F(t)$.*

Proof. If all factors of $v\sigma$ are in normal form, we apply lemma 12. Otherwise ($t \notin G(St(v)\sigma\downarrow)$ and $v\sigma\downarrow \notin F(t)$), choose a constant c, which does not occur in $t, v, \sigma, \mathcal{R}$. Then $v\sigma[t \mapsto u]\downarrow = v\sigma[t \mapsto c]\downarrow[c \mapsto u]\downarrow$. By lemma 11, $v\sigma[t \mapsto c]\downarrow = v\sigma\downarrow[t \mapsto c]$. Then $v\sigma[t \mapsto u]\downarrow = v\sigma\downarrow[t \mapsto c][c \mapsto u]\downarrow = v\sigma\downarrow[t \mapsto u]\downarrow$.

Lemma 13. *For every term v, context ζ and constant c not occuring in $\mathcal{R}$, if $v\downarrow = \zeta[c]$, $\zeta[t] \in F(t)$ and t is in normal form, then $v[c \mapsto t]\downarrow = \zeta[t]$.*

Proof. Since c does not occur in $\mathcal{R}$, we have $v\downarrow = \zeta[c] \implies v[c \mapsto t] \to^* \zeta[t]$. Since $\zeta[t] \in F(t)$ and t is in normal form, we get by hypothesis 4 that $\zeta[t]$ is in normal form. So we conclude $v[c \mapsto t]\downarrow = \zeta[t]$.

Definition 7. *Let v be a term and t be a term in normal form, $c_1, \ldots, c_n$ be the constants of C_{priv} occurring in v, t and $c_t, c_1^p, \ldots, c_n^p$ be constants not belonging to C_{priv} and not occurring in $\mathcal{R}$. Let δ_{c,c^p} be the replacement of every c_i with c_i^p. We let*

$$v^t = v[t \mapsto c_t]\delta_{c,c^p}[c_t \mapsto t]$$

Intuitively, v^t is obtained from v by replacing all constants in v by new constants, except those, which appear in an occurrence of t as a subterm of v. This transformation somehow marks the positions, which do not correspond to occurrences of t as a subterm.

Example 13. For instance, $(a\star b+a+b)^{a\star b} = a\star b+a^p+b^p$, $(a\star b\star c)^{a\star b} = a^p\star b^p\star c^p$ and $((a+a\star b)\star(b+a\star b))^{a\star b} = (a^p+a\star b)\star(b^p+a\star b)$. Note that, if $t \notin St(v)$, then v^t contains only public symbols.

Lemma 14. *If v is in normal form, then v^t is in normal form.*

Proof. v can be obtained from v^t by replacing back each c_i^p with c_i. Since the constants c_i^p do not occur in $\mathcal{R}$, if $l\sigma$ is a redex in v^t, then $l\sigma'$ is a redex in v, where σ' is the substitution σ in which every c_i^p is replaced with c_i.

Lemma 15. *Let u be a term, t be a term in normal form and σ be a substitution. Assume that either all factors of $u\sigma$ are in normal form or that $t \notin G(St(u)\sigma\downarrow)$. Let $v = u\sigma$. Then:*

1. *if $v\downarrow \notin F(t)$, then $v^t\downarrow = v\downarrow^t$*
2. *if $v\downarrow \in F(t)$, then either $v^t\downarrow = v\downarrow$ or $v[t \mapsto c_t]\downarrow = v\downarrow$.*

Proof. **1.** Let $w = v[t \mapsto c_t]\delta_{c,c_p}$. Suppose first that $v\downarrow \notin F(C_{priv})$. Then $v\downarrow \notin F(t) \cup F(C_{priv})$ and hence, by corollary 2,

$$w\downarrow = v\downarrow[t \mapsto c_t]\delta_{c,c_p} \tag{1}$$

Suppose that $w\downarrow = c_t$. By (1), this can happen only if $v\downarrow = t \in F(t)$: a contradiction.

Thus $w\downarrow \neq c_t$. Then, since $F(c_t) = \{c_t\}$ (hypothesis 5), by applying again the corollary 2, we conclude $w[c_t \mapsto t]\downarrow = w\downarrow[c_t \mapsto t]\downarrow = v\downarrow[t \mapsto c_t]\delta_{c,c_p}[c_t \mapsto t]\downarrow = (v\downarrow)^t\downarrow = (v\downarrow)^t$, using the lemma 14 for the last step.

Suppose now that $v\downarrow \in F(C_{priv})$. By hypothesis 5, we get $v\downarrow = c \in C_{priv}$. Therefore, by lemma 13 in conjunction with corollary 2, we get $w\downarrow = c^p$. So we deduce $w[c_t \mapsto t]\downarrow = w\downarrow[c_t \mapsto t]\downarrow = c^p = v\downarrow^t$.

2. By corollary 2, either $v[t \mapsto c_t]\downarrow = v\downarrow$ (and we conclude) or else $v[t \mapsto c_t]\downarrow = \zeta[c_t]$, for some ζ s.t. $v\downarrow = \zeta[t] \in F(t)$. Note that $\zeta[c_t] \notin F(C_{priv})$, hence by applying again the corollary 2 we get: $w\downarrow = \zeta[c_t]\delta_{c,c_p} = \zeta[c_t]$.

Finally, by lemma 13, we get $v^t\downarrow = w[c_t \mapsto t]\downarrow = \zeta[t] = v\downarrow$, and we conclude.

Lemma 16. *Let v be a term and σ be a substitution. We have: $(v\sigma)\delta_{c,c^p} = (v\delta_{c,c^p})(\sigma\delta_{c,c^p})$*

Proof. Immediate by induction and the definition of replacement.

Lemma 17. *Let v be a term, t be a term in normal form and σ be a substitution s.t. $t \notin G(St(v)\sigma)$. Let δ_{c,c_p} be the replacement of all private constants by a corresponding public one. Then $(v\sigma)^t = v\delta_{c,c_p}(\sigma^t)$.*

Proof. We have:

$$
\begin{array}{ll}
(v\sigma)^t = (v\sigma)[t \mapsto c_t]\delta_{c,c^p}[c_t \mapsto t] = & \textit{by definition} \\
(v(\sigma[t \mapsto c_t]))\delta_{c,c^p}[c_t \mapsto t] = & \textit{by lemma 11} \\
(v\delta_{c,c^p}(\sigma[t \mapsto c_t]\delta_{c,c^p}))[c_t \mapsto t] = & \textit{by lemma 16} \\
= v\delta_{c,c^p}(\sigma[t \mapsto c_t]\delta_{c,c^p}[c_t \mapsto t]) = & \textit{by lemma 11} \\
= v\delta_{c,c_p}(\sigma^t) & \textit{by definition}
\end{array}
$$

Lemma 18. *Let ζ be a recipe s.t. $Fact(\zeta) = \{x_1, \ldots, x_k\} \subseteq X$ and $t, v_1, \ldots, v_k$ be terms. Then we have $\zeta[v_1, \ldots, v_k]^t = \zeta[v'_1, \ldots, v'_k]$, where:*

- *$v'_i = v_i^t$, if $v_i \notin F(t)$.*
- *$v'_i \in \{v_i, v_i^t\}$, otherwise.*

Proof. Follows easily by lemma 10, hypothesis 4 and the definition of $(.)^t$:
By definition, $\zeta[v_1, \ldots, v_k]^t = \zeta[v_1, \ldots, v_k][t \mapsto c_t]\delta_{c,c^p}[c_t \mapsto t]$.
By lemma 10, $\zeta[v_1, \ldots, v_k][t \mapsto c_t] = \zeta[v''_1, \ldots, v''_k]$, where:

- $v''_i = v_i[t \mapsto c_t]$ if $v_i \notin F(t)$
- $v''_i \in \{v_i\} \cup \chi[c_t]$ otherwise.

By lemma 10 and hypothesis 4, we have: $\zeta[v''_1, \ldots, v''_k]\delta_{c,c^p} = \zeta[v''_1\delta_{c,c^p}, \ldots, v''_k\delta_{c,c^p}]$.
Finally, we conclude $\zeta[v''_1\delta_{c,c^p}, \ldots, v''_k\delta_{c,c^p}][c_t \mapsto t] = \zeta[v'_1, \ldots, v'_k]$ with:

- $v'_i = v_i^t$, if $v_i \notin F(t)$.
- $v'_i \in \{v_i, v_i^t\}$, otherwise.

Lemma 19. *Let $C = \{T_1 \Vdash x_1, \ldots, T_n \Vdash x_n\} \cup S$ be a satisfiable constraint system and σ be a solution of C. Let $\zeta_1, \ldots, \zeta_n$ be the corresponding proofs of $T_1\sigma \vdash x_1\sigma, \ldots, T_n\sigma \vdash x_n\sigma$. Let t be a term in normal form s.t. $t \notin G(St(C)\sigma\downarrow)$. We have that $\forall i \forall \zeta \in St(\zeta_i) \forall v$ s.t. $\zeta[T_i\sigma]\downarrow = v$:*

1. *$(T_i\sigma)[t \mapsto c] \vdash v[t \mapsto c]$.*
2. *$(T_i\sigma)[t \mapsto c] \vdash v^t$.*

Where c is a public constant.

Proof. We proceed by induction on the pair (i, ζ).

Base case. If $i = 1$ and ζ is a variable, then $v \in T_1\sigma\downarrow = T_1$, since T_1 is ground by origination. Since $t \notin G(St(T_1)\sigma\downarrow) = G(St(T_1))$, we have $T_1\sigma[t \mapsto c] = T_1\sigma$ and $v[t \mapsto c] = v$. Moreover, v^t contains only public symbols and therefore has a trivial proof. Hence we conclude.

Induction step. We prove the two points separately. Note that by monotonicity, origination and induction hypothesis we have: $\forall x \in var(T_i).(T_i\sigma)[t \mapsto c] \vdash x\sigma[t \mapsto c]$ and $(T_i\sigma)[t \mapsto c] \vdash x\sigma^t$.

1.

a. Assume first that ζ is a variable. Thus, there is an $u \in T_i\sigma$ s.t. $u\downarrow = v$. Let $u = u'\sigma$, with $u' \in T_i$. If $v \notin F(t)$ and since $t \notin G(St(u')\sigma\downarrow)$, we deduce by corollary 2 that: $u[t \mapsto c]\downarrow = v[t \mapsto c]$ and we conclude. Otherwise, if $v \in F(t)$,

by lemma 15, either $u[t \mapsto c]\downarrow = u\downarrow = v$ or $u^t\downarrow = u\downarrow = v$. In the first case, by hypothesis 4, we have $v = v[t \mapsto c]$, and we conclude, or else $v = t$ and $v[t \mapsto c] = c$, and we have a trivial proof of $(T_i\sigma)[t \mapsto c] \vdash v[t \mapsto c]$, since c is a public constant.

In the second case, since $t \notin G(St(u')\sigma)$, by lemma 17, we get $u^t = (u'\sigma)^t = u'\delta_{c,c^p}(\sigma^t)$. Finally, since by induction hypothesis we have $(T_i\sigma)[t \mapsto c] \vdash \sigma^t$ and $u'\delta_{c,c^p}$ contains only public constants, we deduce that $(T_i\sigma)[t \mapsto c] \vdash u'\delta_{c,c^p}(\sigma^t)\downarrow = u^t\downarrow = v$, and we conclude as above that $(T_i\sigma)[t \mapsto c] \vdash v[t \mapsto c]$.

b. Now let $\zeta = \zeta_0[\zeta_1, \ldots, \zeta_k], Fact(\zeta) = \{\zeta_1, \ldots, \zeta_k\}, v_1 = \zeta_1[T_i\sigma]\downarrow, \ldots, v_k = \zeta_k[T_i\sigma]\downarrow$.

Suppose first that $v \notin F(t)$. Since $\zeta[v_1, \ldots, v_k]$ is a term with all its factors in normal form, by corollary 2, we have $v[t \mapsto c] = \zeta_0[v_1, \ldots, v_k][t \mapsto c]\downarrow$ (*). By lemma 10, we get $\zeta_0[v_1, \ldots, v_k][t \mapsto c] = \zeta_0[v'_1, \ldots, v'_k]$ (**). Where:

- either $v'_i = v_i[t \mapsto c]$ and we have a proof of $(T_i\sigma)[t \mapsto c] \vdash v'_i$ by induction hypothesis.
- or else $v'_i = \chi[c]$ and we have a trivial proof of $(T_i\sigma)[t \mapsto c] \vdash v'_i$, since all the symbols in $\chi[c]$ are public.
- or else $v'_i = v_i \in F(t)$. Suppose that $v_i \in F(t)\backslash t$. Then, by hypothesis 4, we have $v'_i = v_i = v_i[t \mapsto c]$, so we have a proof of $(T_i\sigma)[t \mapsto c] \vdash v'_i$ by induction hypothesis. We are left with the case $v'_i = v_i = t$. In this case $v_i^t = v_i$, so we have a proof of $(T_i\sigma)[t \mapsto c] \vdash v'_i$ by the second point of the induction hypothesis.

Let us denote by $\zeta'_1, \ldots, \zeta'_k$ these proofs. Hence we deduce by (*) and (**) that we have a proof $\zeta_0[\zeta'_1, \ldots, \zeta'_k]$ of $(T_i\sigma)[t \mapsto c] \vdash v[t \mapsto c]$, when $v \notin F(t)$.

Suppose now that $v \in F(t)$ and let $v' = \zeta_0[v_1, \ldots, v_k]$. Since v' is with all it's factors in normal form, by lemma 15, either $v'[t \mapsto c]\downarrow = v'\downarrow = v$ or else $v'^t\downarrow = v'\downarrow = v$. If we are in the first case, we conclude: the proof of $(T_i\sigma)[t \mapsto c] \vdash v$ is $\zeta_0[\zeta'_1, \ldots, \zeta'_k]$ and $v[t \mapsto c] \in \{v, c\}$ implies $(T_i\sigma)[t \mapsto c] \vdash v[t \mapsto c]$.

Otherwise, by lemma 18, we have $v'^t = \zeta_0[v'_1, \ldots, v'_k]$, where for all i:

- either $v'_i = v_i^t$ and we have a proof of $(T_i\sigma)[t \mapsto c] \vdash v'_i$ by (the second point of) the induction hypothesis.
- or else $v'_i = v_i \in F(t)$. If $v_i \in F(t)\backslash t$, then $v_i = v_i[t \mapsto c]$ and therefore we have a proof of $(T_i\sigma)[t \mapsto c] \vdash v'_i$ by the first point of the induction hypothesis. If $v_i = t$, then $v_i^t = t$ and we have a proof of $(T_i\sigma)[t \mapsto c] \vdash v'_i$ by the second point of the induction hypothesis.

Hence we deduce $(T_i\sigma)[t \mapsto c] \vdash v'^t\downarrow$ and thus $(T_i\sigma)[t \mapsto c] \vdash v$. This concludes the first part of the lemma: $(T_i\sigma)[t \mapsto c] \vdash v[t \mapsto c]$, since $v[t \mapsto c] \in \{v, c\}$.

2.

a. Assume first that ζ is a variable. Then there is an $u \in T_i\sigma$ s.t. $v = u\downarrow$. Let $u = u'\sigma$, with $u' \in T_i$. Since $t \notin G(St(u')\sigma)$, by lemma 17, we get $u^t = (u'\sigma)^t = u'\delta_{c,c^p}(\sigma^t)$. Since, by induction hypothesis, we have $(T_i\sigma)[t \mapsto c] \vdash \sigma^t$ and $u'\delta_{c,c^p}$ contains only public constants, we deduce that $(T_i\sigma)[t \mapsto c] \vdash u'\delta_{c,c^p}(\sigma^t)\downarrow = u^t\downarrow$.

Suppose first that $v \notin F(t)$. Then, since $t \notin G(St(u')\sigma)$, by lemma 15, we have $v^t = u^t\downarrow$, and we conclude: $(T_i\sigma)[t \mapsto c] \vdash v^t$. If $v \in F(t)$, by lemma 15,

either $u[t \mapsto c]\downarrow = v$ or else $u^t\downarrow = v$. In both cases, we have then $(T_i\sigma)[t \mapsto c] \vdash v$. If $v = t$, then $v^t = t$ so we get a proof of $(T_i\sigma)[t \mapsto c] \vdash v^t$. If $v \in F(t)\backslash\{t\}$, then v^t contains only public symbols, so we have a trivial proof of $(T_i\sigma)[t \mapsto c] \vdash v^t$.

b. Now let $\zeta = \zeta_0[\zeta_1, \ldots, \zeta_k], Fact(\zeta) = \{\zeta_1, \ldots, \zeta_k\}, v_1 = \zeta_1[T_i\sigma]\downarrow, \ldots, v_k = \zeta_k[T_i\sigma]\downarrow$.

Let $v' = \zeta_0[v_1, \ldots, v_k]$. First we show that we have a proof of $(T_i\sigma)[t \mapsto c] \vdash v'^t\downarrow$. Indeed, by lemma 18, we have $v'^t = \zeta_0[v'_1, \ldots, v'_k]$, with $v'_i = v_i^t$ or $v'_i \in \{v_i, v_i^t\}$, when $v_i \in F(t)$. Hence, by applying the induction hypothesis as above, we deduce $(T_i\sigma)[t \mapsto c] \vdash v'^t\downarrow$.

Now, if $v \notin F(t)$, we conclude that $(T_i\sigma)[t \mapsto c] \vdash v^t$ by lemma 15, since $v'^t\downarrow = v^t$. If $v \in F(t)\backslash\{t\}$, then v^t contains only public symbols so we have a trivial proof of $(T_i\sigma)[t \mapsto c] \vdash v^t$. If $v = t$ and $v^t = t$, by lemma 15, we have either $v'^t\downarrow = v$ or else $v'[t \mapsto c]\downarrow = v$. So, we conclude, since we have seen that both $(T_i\sigma)[t \mapsto c] \vdash v'^t\downarrow$ and $(T_i\sigma)[t \mapsto c] \vdash v'[t \mapsto c]\downarrow$ (when proving the first point).

This concludes the proof of the lemma.

Lemma 3 *Let $C = \{T_1 \Vdash x_1, \ldots, T_n \Vdash x_n\} \cup S$ be a constraint system. Then C is satisfiable if and only if it has a conservative solution σ': $St(C\sigma'\downarrow) \subseteq F(G(St(C)\sigma'\downarrow))$.*

Proof. Let σ be a solution to C s.t. there is a $t \in St(C\sigma\downarrow)\backslash F(G(St(C)\sigma\downarrow))$. By lemma 11, we have $t \in St(\sigma)$. We show that $\sigma' = \sigma[t \mapsto c]$ is a solution to C.

First, by lemma 19, we get $\forall i.(T_i\sigma)[t \mapsto c] \vdash x_i\sigma[t \mapsto c]$.

Second, by lemma 11, we get $\forall i.(T_i\sigma)[t \mapsto c] = T_i(\sigma[t \mapsto c])$. Therefore, σ' is a solution to the deducibility contraints in C.

Moreover, let $s_1 = s_2$ be an equation in S. Since $s_1\sigma\downarrow = s_2\sigma\downarrow$, by using the lemma 11, we obtain $s_1(\sigma[t \mapsto c])\downarrow = s_1\sigma\downarrow[t \mapsto c] = s_2\sigma\downarrow[t \mapsto c] = s_2(\sigma[t \mapsto c])\downarrow$. Thus σ' is also a solution to S and we conclude.

C Well-moded theories satisfy our hypotheses

We recall the definition of *well-moded* systems of [7]. Assuming a finite ordered set of modes M, for every function symbol f, $m(f) \in M^+$ is a sequence of modes such that $m(f) = m_1, \ldots, m_n \to m$ if the arity of f is n and $m \geq m_i$ for every i. When $f \in \mathcal{F}_{ac}$, $m(f) = m_1^* \to m$ and $m \geq m_1$: all arguments are assigned the same mode m_1. The mode $m(t)$ of a term $t = f(t_1, \ldots, t_n)$ is m if $m(f) = m_1, \ldots, m_n \to m$ (resp. $f \in \mathcal{F}_{ac}$ and $m(f) = m_1^* \to m$). The *Factors* of a term t are the direct semantic subterms of t. They are defined by: $\mathrm{Fact}_{WM}(x) = \{x\}$ if x is a variable and, if $m(f) = m_1, \ldots, m_n \to m$,

$$\mathrm{Fact}_{WM}(f(t_1, \ldots, t_n)) = \{t_i \,|\, t_i \in \mathcal{X} \mathrm{or} m(t_i) \neq m_i\} \cup \bigcup_{j, t_j \notin \mathcal{X}, m(t_j) = m_j} \mathrm{Fact}_{WM}(t_j)$$

Then $\mathtt{Sub}_{WM}(t) \stackrel{\mathrm{def}}{=} \{t\} \cup \mathtt{Sub}_{WM}(\mathrm{Fact}_{WM}(t))$.

A term t is *well-moded* if $\mathrm{Fact}_{WM}(t) \subseteq \mathcal{X}$. A rewrite system $\mathcal{R}$ is well-moded if, for every rule $l \to r$, l, r are well-moded and either r is a variable or $m(l) = m(r)$.

Example 14. Consider the EP example, let $M = \{0, 1\}$, with $1 > 0$ and m is defined by $h : 0 \to 1, exp : 1, 0 \to 1$, $+, \star : 0^* \to 0$, $\bullet : 1^* \to 1$, then the rewrite system is well-moded.

There is no other mode assignment (besides the trivial one, in which there is a single mode) for which the rewrite system is well-moded.

Let $t = (x \star y + x) \bullet exp(x, y) + z$. Following the above mode assignments, $\mathtt{Sub}_{WM}(t) = t \cup \{(x \star y + x) \bullet exp(x, y), x \star y + x, x, y, z\}$.

Lemma 20. *Taking $F(t) = G(t) = \{t\}$, $\mathtt{Fact}_{WM}()$ satisfies our hypotheses.*

Proof.
- Hypothesis 1, $\mathrm{Fact}(u\sigma) \subseteq \mathrm{Fact}(u)\sigma \cup \mathrm{Fact}(\sigma) \cup G(\sigma)$, follows immediately from definitions: each $x \in var(u)$ is either instanciated with a term having the expected mode, and then the factors fall in $\mathtt{Fact}_{WM}(x\sigma)$, or else $x\sigma$ is itself a factor of $u\sigma$.
- Hypothesis 2 follows from lemma 1 of [7].
- Hypothesis 3 follows from lemma 2 of [7].
- Hypothesis 4 is trivially true, since $F(t) = G(t) = \{t\}$.
- Hypothesis 5 follows from the fact that any constant that does not appear in $\mathcal{F}$ is ill-moded, i.e. it does not match any expected mode.
- Finally, the recipe required by hypothesis 6 is the maximal well-moded recipe ζ s.t. $u = \zeta[\mathtt{Fact}_{WM}(u)]$.

We deduce from theorem 2:

Corollary 3. *For local well-moded systems that satisfy the finite variant property, the satisfiability of deducibility constraint systems is reducible to the satisfiability of pure constraint systems.*

D EP satisfies the required hypotheses

Lemma 21 (Hypothesis 1.). *For every substition σ and every term v,*

$$Fact_{EP}(v\sigma) \subseteq Fact_{EP}(v)\sigma \cup Fact_{EP}(\sigma) \cup G(\sigma)$$

Example 15. Let us show when a factor is in $G(\sigma) \cup \mathrm{Fact}_{EP}(\sigma)$:

- $v = exp(x, t)$ and $x\sigma = h(a + b)$. Then we have $a + b \in \mathrm{Fact}_{EP}(v\sigma)$ and $a + b \in G(\sigma)$.
- $v = x \bullet t$ and $x\sigma = h(a \star b)$. Then we have $a \star b \in \mathrm{Fact}_{EP}(v\sigma)$ and $a \star b \in G(\sigma)$.
- $v = x + c$ and $x\sigma = a + b$. Then we have $a, b \in \mathrm{Fact}_{EP}(v\sigma)$ and $a, b \in \mathrm{Fact}_{EP}(x\sigma)$.

Proof. If v is a variable, we are done. If not, we consider the possible cases for $f = top(v)$.

Case $f \in \{+, J_+, e_+\}$**.** Let $\text{Fact}_{EP}(v) = \{v_1, \ldots, v_n\}$. Let $\{v_1, \ldots, v_n\} = \{u_1, \ldots, u_q\} \cup \{w_1, \ldots, w_p\}$, s.t. each w_i is not a variable. Let $\{u_1, \ldots, u_q\} = \{x_1, \ldots, x_n\} \cup \{y_1, \ldots, y_m\}$ s.t. $\forall i. top(x_i\sigma) \notin \{+, J_+, e_+\}$ and $\forall i. top(y_i\sigma) \in \{+, J_+, e_+\}$. Then, by the definition of Fact_{EP}, we have

$$\text{Fact}_{EP}(v\sigma) = \{w_1\sigma, \ldots, w_p\sigma, x_1\sigma, \ldots, x_n\sigma\} \cup \text{Fact}_{EP}(y_1\sigma) \cup \ldots \cup \text{Fact}_{EP}(y_m\sigma)$$

So we have $\text{Fact}_{EP}(v\sigma) \subseteq \text{Fact}_{EP}(v)\sigma \cup \text{Fact}_{EP}(\sigma)$ and we are done.

Case $f \in \{\star, J_\star, e_\star\}$**.** This case is exactly as the previous one.

Case $f = h$**.** Let $\{w_1, \ldots, w_p\} = \text{Fact}_{EP}(v) \cap var(v)$ and $\text{Fact}_{EP}(v\sigma) = F \cup F_1 \cup \ldots \cup F_p$, where F_i are the syntactic subterms of $w_i\sigma$ that are factors of $v\sigma$ and $F = (\text{Fact}_{EP}(v) \backslash \mathcal{X})\sigma$.

By the definition of factors, we have for each i:

- either $F_i = Fact_+(w_i\sigma)$ and then $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma)$
- or else $F_i = Fact_\star(w_i\sigma)$ and then $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma)$
- or else $v = h(w_i)$ and $F_i = \{w_i\sigma\}$.

In all the cases, we have seen that $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma)$, thus we obtain $\text{Fact}_{EP}(v\sigma) \subseteq \text{Fact}_{EP}(v)\sigma \cup \text{Fact}_{EP}(\sigma)$ and we conclude.

Case $f = exp$**.** Let $\{w_1, \ldots, w_p\} = \text{Fact}_{EP}(v) \cap var(v)$ and $\text{Fact}_{EP}(v\sigma) = F \cup F_1 \cup \ldots \cup F_p$, where F_i are the syntactic subterms of w_i that are factors of $v\sigma$ and $F = (\text{Fact}_{EP}(v) \backslash \mathcal{X})\sigma$.

By the definition of factors, we have for each i:

- either $F_i = \text{Fact}_\star(w_i\sigma)$. In this case we have $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma)$;
- or else $F_i = \text{Fact}_{exp}(w_i\sigma)$. In this case, either $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma)$;
- or else $w_i\sigma = h(t), \top(w_i\sigma) = \bullet$ and $F_i = Fact_{exp}(h(t)) = \{t\} \in G(w_i\sigma)$.

In all the cases, we have seen that $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}(w_i\sigma) \cup G(w_i\sigma)$, thus we obtain $\text{Fact}_{EP}(v\sigma) \subseteq \text{Fact}_{EP}(v)\sigma \cup \text{Fact}_{EP}(\sigma) \cup G(\sigma)$ and we conclude.

Case $f \in \{\bullet, J_\bullet, e_\bullet\}$**.** Let $\{w_1, \ldots, w_p\} = \text{Fact}_{EP}(v) \cap var(v)$ and $\text{Fact}(v\sigma) = F \cup F_1 \cup \ldots \cup F_p$, where F_i are the syntactic subterms of w_i that are factors of $v\sigma$ and $F = (\text{Fact}_{EP}(v) \backslash \mathcal{X})\sigma$.

By the definition of factors, we have for each i:

- either $F_i = Fact_+(w_i\sigma)$ and then $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma)$;
- or else $F_i = Fact_\bullet(w_i\sigma)$ and then either $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma)$;
- or else $w_i\sigma = h(t), \top(w_i\sigma) = exp$ and $F_i = \text{Fact}_\bullet(h(t)) = \{t\} \in G(w_i\sigma)$.

In all the cases, we have seen that $F_i \subseteq \{w_i\sigma\} \cup \text{Fact}_{EP}(w_i\sigma) \cup G(w_i\sigma)$, thus we obtain $\text{Fact}_{EP}(v\sigma) \subseteq \text{Fact}_{EP}(v)\sigma \cup \text{Fact}_{EP}(\sigma) \cup G(\sigma)$ and we conclude.

Lemma 22 (Hypotheses 2 and 3 for AG). *Consider $\circ \in \{+, \star, \bullet\}$ and let v be a term with all it's factors in normal form s.t. $\top(v) = \circ$. Assume that only rewrite rules from $R_{AG(\circ)}$ are used in the normalisation of v. Moreover, if $\circ = \bullet$, then $v = \zeta_\bullet[v_1, \ldots, v_k]$, with $\forall i. \top(v_i) \notin \{\bullet, h\}$. Let t be a term in normal form. Then:*

- *either $\top(v\downarrow) = \circ$ and $Fact_{EP}(v\downarrow) \subseteq Fact_{EP}(v)$ (*) and:*
  - *if $v\downarrow \neq t$, then $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$;*
  - *if $v\downarrow = t$, then $v[t \mapsto c]\downarrow = v\downarrow$.*
- *or else $\top(v\downarrow) \neq \circ$ and $v\downarrow \in Fact_{EP}(v)$ (**) and:*

- *if* $v\downarrow \neq t$, *then* $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$;
- *if* $v\downarrow = t$, *then* $v[t \mapsto c]\downarrow = c$.

Proof. Consider the disjoint combination of $R_{AG(\circ)}$ and the free theory of the remaining symbols. Let Fact_d be the notion of factors defined as in [6] for this disjoint combination. Then, by lemma 3 in [6], we have:

- either $\top(v\downarrow) = \circ$ and $\mathrm{Fact}_d(v\downarrow) \subseteq \mathrm{Fact}_d(v)$
- or else $\top(v\downarrow) \neq \circ$ and $v\downarrow \in \mathrm{Fact}_d(v)$.

Moreover, note that, under the assumptions of the lemma, we have $\mathrm{Fact}_d(v) = \mathrm{Fact}_\circ(v) = \mathrm{Fact}_{EP}(v)$ and, when $\top(v\downarrow) = \circ$, $\mathrm{Fact}_d(v\downarrow) = \mathrm{Fact}_\circ(v\downarrow) = \mathrm{Fact}_{EP}(v\downarrow)$. So we deduce (*) or else (**).

For each $u \in \mathrm{Fact}_{EP}(v)$, let $u' = u[t \mapsto c]$. Let $\mathrm{Fact}_{EP}(v) = \mathrm{Fact}_d(v) = \{u_1, \ldots, u_k\}$. Note that, by the definition of replacement , we have $v[t \mapsto c] = v[u_1 \mapsto u'_1] \ldots [u_k \mapsto u'_k] = v\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k}$, where $\delta_{u,u'}$ denotes the replacement with respect to Fact_d.

Suppose first that (*) holds, i.e. $\top(v\downarrow) = \circ$ and $\mathrm{Fact}_{EP}(v\downarrow) \subseteq \{u_1, \ldots, u_k\}$. Then, by lemma 4 in [6], we have $v\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k}\downarrow = v\downarrow\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k}$. If $v\downarrow \neq t$, we have $v\downarrow[t \mapsto c] = v\downarrow\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k}$ and therefore $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$. If $v\downarrow = t$, we have for all $u_i \in \mathrm{Fact}_{EP}(v\downarrow)$, $u'_i = u_i[t \mapsto c] = u_i$, since u_i is in this case smaller than t and can not contain it as a subterm. Therefore we obtain $v[t \mapsto c]\downarrow = v\downarrow\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k} = v\downarrow$.

Suppose now that (**) holds, i.e. $\top(v\downarrow) \neq \circ$ and $v\downarrow = u_i \in \mathrm{Fact}_{EP}(v)$. Let $v' = v \circ \Box$, for some fresh constant $\Box$. Obviously, $v'\downarrow = v\downarrow \circ \Box \notin \mathrm{Fact}_{EP}(v')$. Hence, by lemma 4 in [6], we have $v'\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k}\downarrow = v'\downarrow\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k}$. Therefore, $v\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k}\downarrow \circ \Box = v\downarrow\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k} \circ \Box$ and hence $v[t \mapsto c]\downarrow = v\downarrow\delta_{u_1,u'_1} \ldots \delta_{u_k,u'_k} = u'_i$, for some i. If $u_i \neq t$, we get $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$. Otherwise, we get $v[t \mapsto c]\downarrow = c$ and we conclude.

Lemma 23 (Hypotheses 2 and 3.). *If v is a term whose factors are in normal form and t is a term in normal form, then:*

A. *either* $\top(v\downarrow) = \top(v)$ *and* $Fact_{EP}(v\downarrow) \subseteq Fact_{EP}(v)$ *and:*
  - *if* $v\downarrow \notin F(t)$, *then* $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$;
  - *if* $v\downarrow \in F(t)$, *then* $v[t \mapsto c]\downarrow \in \{v\downarrow, c\}$.

B. *or else* $\top(v\downarrow) \neq \top(v)$ *and* $v\downarrow \in F(Fact(v))$ *and:*
  - *if* $v\downarrow \notin F(t)$, *then* $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$;
  - *if* $v\downarrow = \zeta[t] \in F(t)$, $t \in Fact(v)$, *then* $v[t \mapsto c]\downarrow = \zeta[c]$.

C. *or else* $\top(v) \in \{exp, \bullet\}$, $\top(v\downarrow) = h$ *and* $Fact_{EP}(v\downarrow) \subseteq Fact_{EP}(v)$ *and* $v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$.

Example 16. We have:

A. $v = exp(h(a + b), c)$.
B. $v = exp(h((a + b) \star I(c)), c)$.
B. $v = h(a \star b - c) \bullet h(c)$.
C. $v = exp(h(a \star I(b)), b)$.

C. $v = h(a - b) \bullet h(b)$.

Proof. We do an induction on the pair (length of a normalizing derivation of v, size of v). If v is in normal form, we are done, falling in case **A**, by using hypothesis 4 for the second point. Otherwise, let us distinguish the cases for $\top(v)$:

Case $\top(v) \in \{+, \star\}$**:** this case follows by lemma 22.

Case $\top(v) = h$**:** then $v = h(v')$ and $\top(v') \in \{\bullet, exp, h\}$. Therefore, by definition, $\text{Fact}_{EP}(v) = \{v'\}$. Hence, by hypothesis, v' is in normal form. We deduce then that v is in normal form, to conclude.

Case $\top(v) = exp$**:** then either $v = exp(v_0, \zeta_\star[v_1, \ldots, v_k])$ or $v = h(\zeta_\star[v_1, \ldots, v_k])$. We consider each subcase:

Subcase $v = exp(v_0, \zeta_\star[v_1, \ldots, v_k])$**:** Let v_0' be the term obtained by normalizing v_0 with respect to the rules $exp(exp(x, y), z) \to exp(x, y \star z)$, $exp(h(x), y) \to h(x \star y)$. That is, we have:

- if $v_0 = exp(exp(\ldots, exp(u_0, u_1)\ldots), u_k)$, then $v_0' = exp(u_0, u_1 \star \ldots \star u_k)$
- if $v_0 = exp(exp(\ldots, exp(h(u_1), u_2)\ldots), u_k)$, then $v_0' = h(u_1 \star \ldots \star u_k)$
- otherwise, $v_0' = v_0$.

Note that, by the definition of factors, we have

$$\text{Fact}_{EP}(v) = \text{Fact}_{EP}(exp(v_0', \zeta_\star[v_1, \ldots, v_k])).$$

We distinguish the cases according to the head of v_0':

- Case $v_0' = h(v')$. Then we have $v \to^+ h(v' \star \zeta_\star[v_1, \ldots, v_k])$. Moreover, by definition, $\text{Fact}_{EP}(v) = \text{Fact}_\star(v') \cup Fact_\star(\zeta_\star[v_1, \ldots, v_k]) = \text{Fact}_{EP}(h(v' \star \zeta_\star[v_1, \ldots, v_k]))$ and $\top(v) = \top(h(v' \star \zeta_\star[v_1, \ldots, v_k])) = exp$. Moreover, $v[t \mapsto c] \to h(v'[t \mapsto c] \star \zeta_\star[v_1, \ldots, v_k][t \mapsto c]) = h(v' \star \zeta_\star[v_1, \ldots, v_k])[t \mapsto c]$. Hence we can conclude, by applying the induction hypothesis to $h(v' \star \zeta_\star[v_1, \ldots, v_k])$.
- Case $v_0' = exp(v'', v')$. Then we have $v \to^+ exp(v'', (v' \star \zeta_\star[v_1, \ldots, v_k]))$. Moreover, by the definition of factors, $\text{Fact}_{EP}(v) = \text{Fact}_{exp}(v'') \cup \text{Fact}_\star(v') \cup Fact_\star(\zeta_\star[v_1, \ldots, v_k])) = \text{Fact}_{EP}(exp(v'', (v' \star \zeta_\star[v_1, \ldots, v_k])))$ and $v[t \mapsto c] \to exp(v'', (v' \star \zeta_\star[v_1, \ldots, v_k]))[t \mapsto c]$. Hence we can conclude, by applying the induction hypothesis to $exp(v'', (v' \star \zeta_\star[v_1, \ldots, v_k]))$.
- Case $top(v_0') \notin \{exp, h\}$. Then we have $v\downarrow = exp(v_0', \zeta_\star[v_1, \ldots, v_k]\downarrow)$ or $v\downarrow = v_0'$. So we can conclude by induction hypothesis.

Subcase $v = h(\zeta_\star[v_1, \ldots, v_k])$**:** By applying lemma 22 to $v' = \zeta_\star[v_1, \ldots, v_k]$, we get:

- either $\top(v'\downarrow) = \star$ and $\text{Fact}_{EP}(v'\downarrow) \subseteq \text{Fact}_{EP}(v')$. Then, by definition, we have $\top(v\downarrow) = exp = \top(v)$ and $\text{Fact}_{EP}(v\downarrow) \subseteq \text{Fact}_{EP}(v)$ (**A**). Moreover, by induction hypothesis, in this case we have:
  - $v'\downarrow \neq t \implies v'[t \mapsto c]\downarrow = v'\downarrow[t \mapsto c]$. By definition of factors, we deduce $v\downarrow \notin F(t) \implies v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$.
  - $v'\downarrow = t \implies v'[t \mapsto c]\downarrow = v'\downarrow$. Hence we deduce $v\downarrow \in F(t) \implies v[t \mapsto c]\downarrow = v\downarrow$
- or else $\top(v'\downarrow) \neq \star$ and $v'\downarrow \in \text{Fact}_{EP}(v')$. Then:

- if $\top(v'\downarrow) = +$, then $v\downarrow = h(v'\downarrow) \in F(\text{Fact}_{EP})(v)$ (**B**).
- otherwise, $\top(v\downarrow) = h$, $v\downarrow = h(t)$, with $t \in \text{Fact}_{EP}(v)$ (**C**).

Moreover, by induction hypothesis, in this case we have:

- if $v'\downarrow \neq t$, then $v'[t \mapsto c]\downarrow = v'\downarrow[t \mapsto c]$. We deduce $v\downarrow \notin F(t) \implies v[t \mapsto c]\downarrow = v\downarrow[t \mapsto c]$.
- if $v'\downarrow = t$, then $v'[t \mapsto c]\downarrow = c$. We deduce $v[t \mapsto c]\downarrow = h(c)$, falling in case (**B**) (when $\top(v'\downarrow) = +$) or case **C** (otherwise).

Case $\top(v) = \bullet$**:** Then there are $\zeta_\bullet, \zeta_+, w, w_1, \ldots, w_n, u, u_1, \ldots, u_m$, s.t.:

- $v \to^* \zeta_\bullet[w_1, \ldots, w_n] \bullet h(\zeta_+[u_1, \ldots, u_m])$, using only the rules:

$$h(x) \bullet h(y) \to h(x + y)$$
$$h(x) \bullet h(y) \bullet z \to h(x + y) \bullet z$$
$$J_\bullet(h(x)) \to h(J_+(x))$$
$$J_\bullet(h(x) \bullet y) \to h(J_+(x)) \bullet J_\bullet(y)$$

- $w = \zeta_\bullet[w_1, \ldots, w_n]$ and $w\downarrow = w\downarrow_{AG(\bullet)}$.
- $u = \zeta_+[u_1, \ldots, u_m]$ and $u\downarrow = u\downarrow_{AG(+)}$.
- $\text{Fact}_{EP}(v) = \text{Fact}_{EP}(w \bullet h(u)) = \text{Fact}_\bullet(w) \cup \text{Fact}_+(u)$
- $v[t \mapsto c] \to^* w[t \mapsto c] \bullet h(u[t \mapsto c])$

We distinguish the following cases:

- Case $\top(w\downarrow) = \bullet$ & $\top(u\downarrow) = +$. Then $\top(v\downarrow) = \top(v)$.
  In this case, since $w\downarrow = w\downarrow_{AG(\bullet)}$ and $u\downarrow = u\downarrow_{AG(+)}$, we can assume that $\top(w) = \bullet$ and $\top(u) = +$. Therefore we have $\text{Fact}_{EP}(v) = \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u)$.
  Then, by lemma 22, we have $\text{Fact}_{EP}(w\downarrow) \subseteq \text{Fact}_{EP}(w)$ and $\text{Fact}_{EP}(u\downarrow) \subseteq \text{Fact}_{EP}(u)$. Moreover, by the definition of factors, we have $\text{Fact}_{EP}(v\downarrow) = \text{Fact}_{EP}(w\downarrow) \cup \text{Fact}_{EP}(u\downarrow)$. So we obtain $\text{Fact}_{EP}(v\downarrow) \subseteq \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u) = \text{Fact}_{EP}(v)$ (**A**) and $v\downarrow[t \mapsto c] = w\downarrow[t \mapsto c] \bullet h(u\downarrow[t \mapsto c])$.
  Moreover, we have $w[t \mapsto c]\downarrow \in \{w\downarrow[t \mapsto c], w\downarrow\}$ and $u[t \mapsto c]\downarrow \in \{u\downarrow[t \mapsto c], u\downarrow\}$. Hence we deduce that $v[t \mapsto c]\downarrow \in \{v\downarrow[t \mapsto c], v\downarrow\}$ and we conclude.
- Case $\top(w\downarrow) = \bullet$ & $\top(u\downarrow) \neq +$.
  In this case, we can assume that $\top(w) = \bullet$ and $\top(u) = + \ \vee \ u = u\downarrow$.
  First we assume that $\top(u) = +$. So we have $\text{Fact}_{EP}(v) = \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u)$.
  Then, by lemma 22, we have $\text{Fact}_{EP}(w\downarrow) \subseteq \text{Fact}_{EP}(w)$ and $u\downarrow \in \text{Fact}_{EP}(u)$.
  Suppose first that $w\downarrow \neq e_\bullet$. Then, by the definition of factors, we have $\text{Fact}_{EP}(v\downarrow) = \text{Fact}_{EP}(w\downarrow) \cup \{u\}$. So we obtain $\text{Fact}_{EP}(v\downarrow) \subseteq \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u) = \text{Fact}_{EP}(v)$ (**A**).
  Suppose now that $w\downarrow = e_\bullet$. Thus we have $v\downarrow = h(u\downarrow)$. Either $\top(u\downarrow) = \star$ and then, by the definition of F, we get $v\downarrow \in F(u\downarrow) \subseteq F(\text{Fact}_{EP}(u))$ (**B**). Or else $\top(u\downarrow) \notin \{+, \star\}$ and then, by definition, $\top(v\downarrow) = h$ and $\text{Fact}_{EP}(v\downarrow) = \{u\downarrow\} \subseteq \text{Fact}_{EP}(v)$ (**C**).
  Finally, if $u = u\downarrow$, then $\text{Fact}_{EP}(v\downarrow) = \text{Fact}_{EP}(w\downarrow) \cup \{u\downarrow\}$ and $\text{Fact}_{EP}(v) = \text{Fact}_{EP}(w) \cup \{u\}$, and the result follows simmilarly.
  Moreover, we have $w[t \mapsto c]\downarrow \in \{w\downarrow[t \mapsto c], w\downarrow\}$ and $u[t \mapsto c]\downarrow \in \{u\downarrow[t \mapsto c], c\}$. Hence, by the definition of factors for $v\downarrow$, we deduce that $v[t \mapsto c]\downarrow \in \{v\downarrow[t \mapsto c], v\downarrow, h(c)\}$ and we conclude.

– Case $\top(w\downarrow) \neq \bullet$ & $\top(u\downarrow) = +$.
In this case, we can assume that $\top(w) = \bullet \vee w = w\downarrow$ and $\top(u) = +$. First we assume that $\top(w) = \bullet$. So we have $\text{Fact}_{EP}(v) = \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u)$.
Then, by lemma 22, we have $w\downarrow \in \text{Fact}_{EP}(w)$ and $\text{Fact}_{EP}(u\downarrow) \subseteq \text{Fact}_{EP}(u)$. Moreover, by the definition of factors, we have $\text{Fact}_{EP}(v\downarrow) = \{w\downarrow\} \cup \text{Fact}_{EP}(u\downarrow)$. So we obtain $\text{Fact}_{EP}(v\downarrow) \subseteq \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u) = \text{Fact}_{EP}(v)$, when $u\downarrow \neq e_+$ (**A**). If $u\downarrow = e_+$, we get $v\downarrow = w\downarrow \in \text{Fact}_E P(v)$ (**B**).
Finally, if $w = w\downarrow$, then $\text{Fact}_{EP}(v\downarrow) = \{w\downarrow\} \cup \text{Fact}_{EP}(u\downarrow)$ and $\text{Fact}_{EP}(v) = \{w\} \cup \text{Fact}_{EP}(u)$, and the result follows simmilarly.
Moreover, we have $w[t \mapsto c]\downarrow \in \{w\downarrow[t \mapsto c], c\}$ and $u[t \mapsto c]\downarrow \in \{u\downarrow[t \mapsto c], u\downarrow\}$. Hence, by the definition of factors for $v\downarrow$, we deduce that $v[t \mapsto c]\downarrow \in \{v\downarrow[t \mapsto c], v\downarrow, c\}$ and we conclude.
– Case $\top(w\downarrow) \neq \bullet$ & $\top(u\downarrow) \neq +$.
In this case, we can assume that $\top(w) = \bullet \vee w = w\downarrow$ and $\top(u) = + \vee u = u\downarrow$. First we assume that $\top(w) = \bullet$ and $\top(u) = +$. So we have $\text{Fact}_{EP}(v) = \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u)$.
Then, by lemma 22, we have $w\downarrow \in \text{Fact}_{EP}(w)$ and $u\downarrow \in \text{Fact}_{EP}(u)$. Moreover, by the definition of factors, we have $\text{Fact}_{EP}(v\downarrow) = \{w\downarrow\} \cup u\downarrow$. So we obtain $\text{Fact}_{EP}(v\downarrow) \subseteq \text{Fact}_{EP}(w) \cup \text{Fact}_{EP}(u) = \text{Fact}_{EP}(v)$ (**A.**).
Finally, if $u = u\downarrow$ or $w = w\downarrow$, the result follows simmilarly.
Moreover, we have $w[t \mapsto c]\downarrow \in \{w\downarrow[t \mapsto c], c\}$ and $u[t \mapsto c]\downarrow \in \{u\downarrow[t \mapsto c], c\}$. Hence, by the definition of factors for $v\downarrow$, we deduce that $v[t \mapsto c]\downarrow \in \{v\downarrow[t \mapsto c], v\downarrow, c, h(c)\}$ and we conclude.

Lemma 24. *Let u be a term with all it's factors in normal form s.t. $\top(u\downarrow) \neq \top(u)$. Let $\overline{F}(T) = T \cup h(T)$. Then we have $u\downarrow \in \overline{F}(Fact(u))$.*

Proof. By lemma 23, we get:

– either $u\downarrow \in F(Fact(u))$. In this case we are done, since $F(T) \subseteq \overline{F}(T)$.
– or else $\top(u\downarrow) = h$ and $Fact(u\downarrow) \subseteq Fact(u)$. In this case, by the definition of $\top$ and factors, we have $u\downarrow = h(t)$ and $\text{Fact}_{EP}(u\downarrow) = t$. Therefore we conclude $u\downarrow \in \overline{F}(Fact(u))$

Lemma 25 (Hypothesis 4). *for any set of terms T, $F(F(T)) = F(T)$ and for any term t, $Fact(F(t)) = Fact(G(t)) = Fact(t)$ and, if t is in normal form, then every $u \in F(T)$ is in normal form.*

Proof. Recall that $F(t) = \{t, h(t)\}$, if $top(t) \in \{\star, J_\star, +, J_+\}$ and G is the inverse of F. For proving the first equality, it suffices to note that $F(t)\backslash t \subseteq \{h(t)\}$ and $F(h(t)) = \{h(t)\}$.

For proving the second, it suffices to check in the definition of factors that, if $top(t) \in \{+, J_+, \star, J_\star\}$ then $Fact(h(t)) = Fact(t)$.

Checking that $F(t)$ is in normal form, when t is, is also immediate by inspecting $\mathcal{R}_{EP}$.

Lemma 26 (Hypothesis 5). *For any constant c, $F(c) = \{c\}$ and if c does not occur in $\mathcal{R}$, then $\forall \zeta. c \in St(\zeta[c])$.*

Proof. These two points follow immediately from the definitions.

Lemma 27 (Hypothesis 6). *For every term u s.t. $Fact(u) = \{u_1, \ldots, u_k\}$, there is a recipe $\zeta[x_1, \ldots, x_k]$ s.t. $Fact(\zeta) = \{x_1, \ldots, x_k\}$ and $u = \zeta\{x_1 \mapsto u_1, \ldots, x_k \mapsto u_k\}$.*

Proof. We consider the possible cases for $top(u)$ and check the definition of factors:

- If $top(u) \in \{+, J_+, e_+\}$, then $\mathrm{Fact}(u) = Fact_+(u)$: there is a $\zeta_+[x_1, \ldots, x_k]$ s.t. $u = \zeta[u_1, \ldots, u_k]$ and $Fact(\zeta_+) = Fact_+(\zeta_+) = \{x_1, \ldots, x_k\}$.
- If $top(u) \in \{\star, J_\star, e_\star\}$, we conclude as above.
- If $u = h(u')$, then:
  - If $top(u') \in \{+, J_+, e_+, \star, J_\star, e_\star\}$, then $\mathrm{Fact}(u) = \mathrm{Fact}(u')$. Hence $\zeta = h(\zeta')$, where ζ' is the corresponding recipe for u'.
  - otherwise, $\mathrm{Fact}(u) = \{u'\}$. Hence $\zeta = h(x)$.
- If $u = exp(t, v)$, let $v = \zeta_\star[v_1, \ldots, v_n]$. Then $\mathrm{Fact}(u) = F_t \cup Fact_\star(v)$, where:
  - if $t = h(t')$, then $F_t = Fact_\star(t')$. Let $t' = \zeta'_\star[t_1, \ldots, t_m]$. Then $\zeta = exp(\zeta'_\star, \zeta_\star)$.
  - if $t = exp(t_1, t_2)$, then $F_t = \mathrm{Fact}(t)$. Let ζ' be the corresponding context for t'. Then $\zeta = exp(\zeta', \zeta_\star)$.
  - $F_u = \{u\}$, otherwise. Then $\zeta = exp(x, \zeta_\star)$.
- If $top(t) \in \{\bullet, J_\bullet, e_\bullet\}$ and $t = C_\bullet[t_1, \ldots, t_m, h(t'_1), \ldots, h(t'_n)], top(t_i) \notin \{\bullet, J_\bullet, e_\bullet, h\}$, then $\mathrm{Fact}(t) = \{t_1, \ldots, t_m\} \cup Fact_+(t'_1) \cup \ldots \cup Fact_+(t'_n)$. Then we have $\zeta = C_\bullet[x_1, \ldots, x_m, h(\zeta_1^+), \ldots, h(\zeta_n^+)]$, where ζ_i^+ are the recipes corresponding to the definition of $Fact_+$.

E Proof of theorem 2

Theorem 2 *Assume that the notions of factors, subterms and the functions F, G satisfy the hypotheses of the previous section. Assume moreover that the proof system is local and that the rewrite system has the finite variant property.*

Then the satisfiability of deducibility constraint systems is reducible to the satisfiability of pure constraint systems.

Proof. Let C be the constraint system and σ be a solution of C. By lemma 3, C has a solution σ_1 such that $\mathrm{St}(C\sigma_1\downarrow) \subseteq F(G(\mathrm{St}(C)\sigma_1\downarrow))$.

Using the finite variant property, we get rid of the normalization in this inclusion: there is a variant $C\theta_1\downarrow$ such that $\mathrm{St}(C)\sigma_1\downarrow \subseteq \mathrm{St}(C\theta_1\downarrow)\sigma_2$ for some substitution σ_2 such that $\sigma_1 = \theta_1\sigma_2$. Let $CS_1 = C\theta_1\downarrow$. We have obtained $G(\mathrm{St}(C)\sigma_1\downarrow) \subseteq G(\mathrm{St}(CS_1)\sigma_2)$. We may also take σ_2 out of G by further instanciating the variables of CS_1 using a substitution θ_2 whose range is the fixed finite set of interface contexts: there is some $CS_2 = CS_1\theta_2$ such that $G(\mathrm{St}(CS_1)\sigma_2) \subseteq G(\mathrm{St}(CS_2))\sigma_3$. Now, we let $\overline{F}(u)$ be the "upper bound" of $F(u)$:

$\overline{F}(u) = \{C_1(u), \ldots, C_n(u)\}$. All the previous inclusions give us : $\mathrm{St}(C\sigma_1\downarrow) \subseteq F(G(\mathrm{St}(CS_2))\sigma_3) \subseteq \overline{F}(G(\mathrm{St}(CS_2)))\sigma_3$.

By locality, the deducible terms are in $F_1(\mathrm{St}(C\sigma_1\downarrow)) \subseteq F_1(\overline{F}(G(\mathrm{St}(CS_2)))\sigma_3)$. As before, by further instanciating the variables, we may take F_1 out of σ_3, to get $F_1(\mathrm{St}(C\sigma_1\downarrow)) \subseteq \overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3)))\sigma_4$, for some computable θ_3 and $CS_3 = CS_2\theta_3$.

Then, we non-deterministically choose $\theta_1, \theta_2, \theta_3$, add to the equational part of C the equations $x = x\theta_1\theta_2\theta_3$ for each variable of C and then non-deterministically construct a system C' as follows:

- Let $C_0 = \{\} \cup S$.
- For each $1 \leq i \leq n$:
- Let $C_i = \{V_{1,1} \Vdash z_{1,1}, \ldots, V_{i,1+k_i} \Vdash z_{i,1+k_i}\} \cup S_i$
- Guess a (possibly empty) sequence $v_{i+1,1}, \ldots, v_{i+1,k_{i+1}} \in \overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3)))$ of distinct terms. Let $S_{i+1} = S_i \cup \{z_{i+1,1} = v_{i+1,1}, \ldots, z_{i+1,k_{i+1}} = v_{i+1,k_{i+1}}\}$ where $z_{i+1,1}, \ldots, z_{i+1,k_{i+1}}$ are new variables.
- Let $V_{i+1,1} = V_{i,1+k_i} \cup T_{i+1}$ and add to C_i the following constraints:

$$\begin{array}{ll} V_{i+1,1} & \Vdash\ z_{i+1,1} \\ V_{i+1,1}, z_{i+1,1} & \Vdash\ z_{i+1,2} \\ & \ldots \\ V_{i+1,1}, z_{i+1,1}, \ldots, z_{i+1,k_{i+1}-1} & \Vdash\ z_{i+1,k_{i+1}} \\ V_{i+1,1}, z_{i+1,1}, \ldots, z_{i+1,k_{i+1}-1}, z_{i+1,k_{i+1}} & \Vdash\ x_{i+1} \end{array}$$

We let $V_{i+1,j+1} = V_{i+1,1} \cup \{z_{i+1,1}, \ldots, z_{i+1,j}\}$ and $z_{i+1,1+k_{i+1}} = x_{i+1}$

The resulting systems still satisfy the monotonicity and origination properties.

We prove now that C has a solution iff one of the constraints C' obtained by the above non-deterministic algorithm has a pure solution $\sigma \uplus \theta$ where the domain of θ is the set of new variables introduced by the algorithm.

Correctness. Suppose that some C' has a pure solution σ'. Then $\sigma = \sigma'|_{Var(C)}$ is a solution of C: by induction on i, any solution θ of C_i is a solution of $\{T_1 \Vdash x_1, \ldots T_i \Vdash x_i\}$.

Completeness. Let σ be a solution of C. As already explained $\sigma = \theta_1\theta_2\theta_3\sigma_4$ such that for every $T_i \Vdash x_i \in C$, there is a recipe ζ such that $\zeta[T_i]\sigma\downarrow = x_i\sigma$ and for every $\zeta' \in \mathrm{St}(\zeta)$, $\zeta'[T_i]\sigma\downarrow \in \overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3)))\sigma_4$.

We construct then, by induction on ζ, a sequence $S_\zeta = \{t_1, \ldots, t_n\}$ of terms in $\overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3)))$ such that $t_n\sigma_4 = \zeta[T_i]\sigma\downarrow$ and, for every $j = 1, ..., n$, there is a pure recipe ζ_j such that $\zeta_j[T_i\sigma_3, t_1\sigma_4, \ldots, t_{j-1}\sigma_4]\downarrow = t_j\sigma_4$.

If ζ is pure, then the sequence consists of a single element v_ζ such that $v_\zeta \in \overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3)))$ and $v_\zeta\sigma_4 = \zeta[T_i]\sigma\downarrow$. Otherwise, let $\zeta = \zeta'[\zeta_1, \ldots, \zeta_n]$ where $\mathrm{Fact}(\zeta) = \{\zeta_1, \ldots, \zeta_n\}$. Then ζ' is pure and there is a term $t \in \overline{F_1}(\overline{F}(G(\mathrm{St}(CS_3)))$ such that $t\sigma_4 = \zeta'[\zeta_1[T_i]\sigma, \ldots, \zeta_n[T_i]\sigma]\downarrow$. By induction hypothesis, there are sequences of terms $S_{\zeta_1}, \ldots, S_{\zeta_n}$ satisfying the desired properties. We let then S_ζ

be $S_{\zeta_1} \cdot \cdots \cdot S_{\zeta_n} \cdot t$. S_ζ satisfies the desired properties: this is a direct consequence of the induction hypothesis or of the construction of t. For instance, $t\sigma_4 = \zeta'[t_1\sigma_4, \ldots, t_n\sigma_4]\downarrow$ if $t_1\sigma_4 = \zeta_1[T_i]\sigma\downarrow, \ldots t_n\sigma_4 = \zeta_n[T_i]\sigma\downarrow$.

Next, we simply remove doubles from the sequence S_ζ and we get the intermediate terms $v_{i,j}$.

F Solving pure systems

F.1 Reduction to three recipe types

Lemma 5 *In case of EP, any pure and normalized recipe has one of the following forms:*

$$\begin{array}{ll} \zeta_+ : \zeta = \zeta_+[x_1, \ldots, x_n] & \zeta_\star : \zeta = \zeta_\star[x_1, \ldots, x_n] \\ \zeta_h^\star : \zeta = h(\zeta_\star[x_1, \ldots, x_n]) & \zeta_h^+ : \zeta = h(\zeta_+[x_1, \ldots, x_n]) \\ \zeta_{exp}^\star : \zeta = exp(y_0, \zeta_\star[x_1, \ldots, x_n]) & \zeta_\bullet^+ : \zeta = \zeta_\bullet[x_1, \ldots, x_n] \bullet h(\zeta_+[y_1, \ldots, y_m]), n \geq 1 \end{array}$$

Where $\zeta_\circ$ is a recipe in $\mathcal{T}(\{\circ, J_\circ, e_\circ\}, \mathcal{X})$, for $\circ \in \{+, \star, \bullet\}$.

Proof. By the definition of factors, it is clear that these recipes are pure.

Now let ζ be a recipe in normal form that does not fall into one of the cases above. In the following, u is a non-variable term. We have:

Case $top(\zeta) \in \{+, J_+\}$**.** Then $\zeta = \zeta_+[u]$, with $top(u) \notin \{+, J_+, e_+\}$. Hence $u \in St(\zeta)$ and ζ is not pure.

Case $top(\zeta) \in \{\star, J_\star\}$**.** Then $\zeta = \zeta_\star[u]$, with $top(u) \notin \{\star, J_\star, e_\star\}$. Hence $u \in St(\zeta)$ and ζ is not pure.

Case $top(\zeta) = h$**.** Since ζ is not of the type $\zeta_h^\star, \zeta_h^+$ or $\zeta_\bullet^+$, we must have either $\zeta = h(\zeta_+[u]), top(u) \notin \{+, J_+, e_+\}$ or $\zeta = h(\zeta_\star[u]), top(u) \notin \{\star, J_\star, e_\star\}$ or $\zeta = h(u), top(u) \notin \{+, J_+, e_+, \star, J_\star, e_\star\}$. In all these cases, $u \in St(\zeta)$ and hence ζ is not pure.

Case $top(\zeta) = exp$**.** Let $\zeta = exp(\zeta_1, \zeta_2)$. Since ζ is in normal form, we have $top(\zeta_1) \notin \{exp, h\}$. Hence, if ζ_1 is not a variable, ζ is not pure. Moreover, if ζ_2 is not a variable and $\zeta_2 = \zeta_\star[u], top(u) \notin \{\star, J_\star, e_\star\}$, then $u \in St(\zeta)$ and hence ζ is not pure.

Case $top(\zeta) = \bullet$**.** Then either $\zeta = \zeta_\bullet[u], top(u) \notin \{\bullet, J_\bullet, e_\bullet, h\}$ or else $\zeta = \zeta_\bullet[h(\zeta_+[u])], top(u) \notin \{+, J_+, e_+\}$. Then $u \in St(\zeta)$ and hence ζ is not pure.

Lemma 6 *The transformations $C \to \bigvee C_j$ of section 6.1 are correct and complete: the resulting constraints C_j are pure and satisfy monotonicity and origination. Any solution of some C_j is a solution of C and any solution of C is a solution of some C_j.*

Proof. It is enough to prove the lemma when only one constraint is transformed: the general case results by iteration. Also, we will consider only the case $\zeta_{exp}^\star$: the other two cases are similar. First note that, by construction, if $C \to C'$, then C' satisfies monotonicity and origination.

Let $C = \{T_1 \Vdash x_1, \ldots, T_i \Vdash_{\zeta^\star_{exp}} x_i, , T_{i+1} \Vdash x_{i+1}, \ldots, T_n \Vdash x_n\} \cup S$ be s.t. the i-th constraint is transformed. Then $C' = \{T_1 \Vdash x_1, \ldots, T_i \Vdash_\star y_i, T_{i+1}[x_i \mapsto exp(u, y_i)] \Vdash x_{i+1}, \ldots, T_n[x_i \mapsto exp(u, y_i)] \Vdash x_n\} \cup S \cup \{x_i = exp(u, y_i)\}$, for some $u \in T_i$.

Correctness: let $\sigma' = \sigma \cup \{y_i \mapsto t\}$ be a pure solution to C'. To show that σ is a pure solution to C, it is sufficient to show that:

- $T_i\sigma \vdash_{\zeta^\star_{exp}} x_i\sigma$;
- for all $j > i$, $T_j[x_i \mapsto exp(u, y_i)]\sigma'{\downarrow} = T_j\sigma{\downarrow}$;

Note that $x_i\sigma' = exp(u\sigma', y_i\sigma')$ and $y_i\sigma' = \zeta_\star[T_i\sigma']{\downarrow}$. By construction, $u\sigma' = u\sigma$ and $T_i\sigma' = T_i\sigma$. Hence we have $x_i\sigma' = x_i\sigma = exp(u\sigma, y_i\sigma')$ and $y_i\sigma' = \zeta_\star[T_i\sigma]{\downarrow}$. Therefore $x_i\sigma = exp(u\sigma, \zeta_\star[T_i\sigma]){\downarrow}$ and $T_i\sigma \vdash_{\zeta^\star_{exp}} x_i\sigma$.

The second point follows by $exp(u, y_i)\sigma'{\downarrow} = x_i\sigma' = x_i\sigma$ and the convergence of the rewriting system.

Completness: let σ be a pure solution to C. We show that there is a C' and a pure solution σ' to C'. Indeed, since $T_i\sigma \vdash_{\zeta^\star_{exp}} x_i\sigma$, there is a $u \in T_i$ and a (pure) $\zeta_\star$ s.t. $x_i\sigma = exp(u\sigma, \zeta_\star[T_i\sigma]){\downarrow}$. Let us choose this u to construct C'. It is easy to show, following the same lines as above, that $\sigma' = \sigma \cup \{y_i \mapsto \zeta_\star[T_i\sigma]{\downarrow}\}$ is a pure solution to C'. This way we conclude.

F.2 Stabilizing the root symbol: a preliminary lemma

We start with a preparation lemma, which precises, in the case of EP, in which situations we need the function G when computing the factors of an instance.

Lemma 28. *For any pure and linear context $\zeta(x_1, \ldots, x_k)$ and substitution σ, we have $Fact_{EP}(\zeta\sigma) = F_1 \cup \ldots \cup F_k$, where for all i:*

- *either $F_i = \{x_i\sigma\}$;*
- *or else $F_i = Fact_{EP}(x_i\sigma)$;*
- *or else $F_i = G(x_i\sigma)\backslash\{x_i\sigma\}$*

Moreover,

- *$F_i = G(x_i\sigma)\backslash\{x_i\sigma\}$, only if:*
  - *either $\zeta = exp(\ldots exp(exp(x_i, \chi^1_\star), \chi^2_\star), \ldots), \chi^n_\star)$, $x_i\sigma = h(t)$ and $\top(t) = +$.*
  - *or else $\zeta = x_i \bullet \zeta'$, $x_i\sigma = h(t)$ and $\top(t) = \star$.*

  In both cases, we have $F_i = \{t\}$.
- *Let σ and θ be two substitutions s.t. $\top(x_i\sigma) = \top(x_i\theta)$. Then $F_i(\sigma) = \{x_i\sigma\} \implies F_i(\theta) \subseteq G(x_i\theta)$.*

Proof. We consider all the possible pure contexts:

ζ_+: $\zeta = \zeta_+(x_1, \ldots, x_n)$. Let $\{x_1, \ldots, x_n\} = \{y_1, \ldots, y_k\} \cup \{z_1, \ldots, z_m\}$ s.t. for all i, $\top(y_i\sigma) \neq +$ and $\top(z_i\sigma) = +$. Then, by the definition of Fact_{EP}, we have

$$\mathrm{Fact}_{EP}(\zeta\sigma) = \{y_1\sigma, \ldots, y_k\sigma\} \cup \mathrm{Fact}_{EP}(z_1\sigma) \cup \ldots \cup \mathrm{Fact}(z_m\sigma)$$

Thus forall i, $F_i = Fact(x_i\sigma)$ or $F_i = \{x_i\sigma\}$ and we deduce also the second point of the lemma, since F_i depends directly on the top of $x_i\sigma$: $\top(x_i\sigma) = \top(x_i\theta)$ & $F_i(\sigma) = \{x_i\sigma\} \implies F_i(\theta) = \{x_i\theta\}$.

$\zeta_\star$: $\zeta = \zeta_\star[x_1, \ldots, x_n]$. This case is simmilar to the previous one.

$\zeta_h^\star$: $\zeta = h(\zeta_\star[x_1, \ldots, x_n])$. This case is simmilar, except when $\zeta_\star$ is a variable: $\zeta = h(x)$. Then, by the definition of factors, we have $\mathrm{Fact}_{EP}(\zeta\sigma) \subseteq \{x\sigma\} \cup \mathrm{Fact}_{EP}(x\sigma)$. Moreover, the second part of the lemma also follows easily, since $F_x = \{x\sigma\}$ iff $\top(x\sigma) \notin \{+, \star\}$.

$\zeta_{exp}^\star$: $\zeta = exp(\zeta'[x_1, \ldots, x_n], \zeta_\star[y_1, \ldots, y_m])$. Note that, since ζ is pure, we must have $\top(\zeta') = exp$ or $\zeta' \in \{x, h(x)\}$, with $x \in \mathcal{X}$.

By the definition of factors, we have: $\mathrm{Fact}_{EP}(\zeta\sigma) \subseteq F' \cup F_{y_1} \cup \ldots \cup F_{y_m}$, where for all i, either $\top(y_i\sigma) \neq \star$ and $F_{y_i} = \{y_i\sigma\}$ or else $\top(y_i\sigma) = \star$ and $F_{y_i} = \mathrm{Fact}_{EP}(y_i\sigma)$. For F' we have the following cases:

- either (base case) $\zeta' = x$ and then $F' = F_x$, with:
  - $F_x = \mathrm{Fact}_{EP}(x\sigma)$, when $\top(x\sigma) \in \{exp, h\}$
  - $F_x = \{x\sigma\}$, when $\top(x\sigma) \in \{+, \star, \bullet\}$ and $top(x) \neq h$
  - $F_x = G(x\sigma)\backslash\{x\sigma\}$, when $\top(x\sigma) = \bullet$ and $top(x) = h$.

  Note that $\top(x\sigma) = \top(x\theta)$ & $F_x(\sigma) = \{x\sigma\} \implies F_x(\theta) \subseteq G(x\theta)$. Moreover, $F_x = G(x\sigma)\backslash\{x\sigma\}$, only when $\zeta = exp(x, \zeta_\star[y_1, \ldots, y_n])$, $x\sigma = h(t)$ and $\top(t) = +$.
- or else (base case) $\zeta' = h(x)$ and then $F' = F_x$, with $F_x = Fact_\star(x\sigma) \subseteq \{x\sigma\} \cup \mathrm{Fact}_{EP}x\sigma$.
- or else (induction step) $\top(\zeta') = exp$, $F' = \mathrm{Fact}_{EP}(\zeta'\sigma) = F_{x_1} \cup \ldots \cup F_{x_n}$ and the property for $F_{x_1}, \ldots, F_{x_n}$ follows by induction hypothesis.

$\zeta_\bullet^+$: $\zeta = \zeta_\bullet[x_1, \ldots, x_n] \bullet h(\zeta_+[y_1, \ldots, y_m])$.

Let $w \in \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$. By the definition of factors, we have:

- either $F_w = Fact_+(w\sigma)$ and then $F_w \subseteq \{w\sigma\} \cup \mathrm{Fact}_{EP}(w\sigma)$;
- or else $\top(w\sigma) = \bullet$ and $F_w = \mathrm{Fact}_{EP}(w\sigma)$;
- or else $\top(w\sigma) \notin \{\bullet, h\}$ and $F_w = \{w\sigma\}$;
- or else $w\sigma = h(w')$, $\top(w\sigma) \neq \bullet$ and $F_w = \mathrm{Fact}_+(w') = \{w'\}$. If $\top(w) = h$ we have $\mathrm{Fact}_{EP}(h(w')) = \{w'\}$ and thus $F_w = \mathrm{Fact}_{EP}(w\sigma)$. Otherwise, $\top(w) = exp$ and $w' \in G(w\sigma)$, thus we have $F_w \subseteq G(w\sigma)$.

In all the cases, we have seen that $F_w \subseteq \{w\sigma\} \cup \mathrm{Fact}(w\sigma) \cup G(w\sigma)$. Moreover, note that:

- $F_w = G(w\sigma)\backslash\{w\sigma\}$, only if $\zeta = w \bullet \zeta'$, $w\sigma = h(t)$ and $\top(t) = \star$.
- $\top(w\sigma) = \top(w\theta)$ & $F_w(\sigma) = \{w\sigma\} \implies F_w(\theta) \subseteq G(w\theta)$

Hence we conclude the proof of the lemma.

F.3 Stabilizing the root symbol: proof of lemma 7

We show here first a remarquable property of $\mathcal{R}_{EP}$: roughly, the root symbol can change by normalization, only if the term can be reduced to one of its subterms, possibly adding or removing interface symbols.

We will consider the following functions: $\overline{F}(T) = T \cup \{h(t) \mid t \in T\}$, $\overline{G}(T) = T \cup \{t \mid h(t) \in T\}$ and $\mathcal{D}(T) = \overline{F}(T) \cup \overline{G}(T)$. Note that $F(T) \subseteq \overline{F}(T)$ and $G \subseteq \overline{G}(T)$.

Lemma 29. *For every term u s.t. $\top(u\downarrow) \neq \top(u)$, we have $u\downarrow \in \overline{F}(G(St_<(u))\downarrow)$.*

Proof. We proceed by induction on the size of u. If u is a constant or in normal form, the lemma is trivially true, since $\top(u\downarrow) = \top(u)$.

Otherwise, let $u = \zeta(u_1, \ldots, u_k), Fact(u) = \{u_1, \ldots, u_k\}$ and $v = \zeta[u_1\downarrow, \ldots, u_k\downarrow]$. Note that either $\top(v) = \top(u)$ or else $u = h(u')$ and $\top(u') \neq \top(u'\downarrow) \in \{+, \star\}$. In the second case, by induction hypothesis, we have $u'\downarrow \in \overline{F}(G(St_<(u'))\downarrow)$. Since $\top(u'\downarrow) \in \{+, \star\}$, we get $u'\downarrow \in G(St_<(u'))\downarrow$ and hence we conclude $u\downarrow \in \overline{F}(G(St_<(u))\downarrow)$, by using also $St_<(u') \subseteq St_<(u)$.

Let us assume now that $\top(v) = \top(u)$. Since, by convergence, $v\downarrow = u\downarrow$, we have $\top(v\downarrow) \neq \top(v)$. Therefore, by lemma 24, we get $u\downarrow = v\downarrow \in \overline{F}(Fact(v)) = \overline{F}(Fact(\zeta(u_1\downarrow, \ldots, u_k\downarrow)))$. By lemma 28, one of the following holds:

- $\exists i. u\downarrow \in \overline{F}(u_i\downarrow)$.
- $\exists i. u\downarrow \in \overline{F}(\mathrm{Fact}(u_i\downarrow))$ and $\top(u_i\downarrow) \neq \top(u_i)$ (otherwise, by the last point of lemma 28 and since $u_i \in \mathrm{Fact}(\zeta(u_1, \ldots, u_k))$, we can descend at most in $G(u_i\downarrow)$, not in $\mathrm{Fact}(u_i\sigma\downarrow)$)
- $v = exp(exp(\ldots exp(h(u\downarrow), v_1), \ldots, v_m)$ and $\top(u\downarrow) = +$ or $v = h(u\downarrow) \bullet v'$ and $\top(u\downarrow) = \star$.
- $v = exp(exp(\ldots exp(h(u'), v_1), \ldots, v_m)$ or $v = h(u') \bullet v'$, with $u\downarrow = h(u')$ and $u' \in G(u_i\sigma\downarrow) \backslash \{u_i\sigma\downarrow\}$, for some i.

If we are in the first case, we conclude. If we are in the fourth case, we get $u\downarrow = u_i\downarrow$, for some i, so we also conclude.

Suppose now that we are in the second case. Then, by applying the induction hypothesis for u_i, we get $u_i\downarrow \in \overline{F}(G(St_<(u_i))\downarrow)$. Let w be a term in $G(St_<(u_i))$ s.t. $u_i\downarrow \in \overline{F}(w\downarrow) = \{w\downarrow, h(w\downarrow)\}$. Let $w' \in \{w, h(w)\}$ be s.t. $w'\downarrow = u_i\downarrow$ and consider the term $u' = \zeta(u_1, \ldots, u_{i-1}, w', u_{i+1}, \ldots, u_k)$.

Let us show that we may apply induction hypothesis to u'. Indeed, u' is not smaller than u only if $u_i = h(w)$. In this case, we can have $\top(u_i\downarrow) \neq \top(u_i)$ only if $\top(w\downarrow) \neq \top(w)$ and one of $\top(w\downarrow), \top(w)$ is in $\{+, \star\}$. Moreover, since $w \in St_<(u_i)$, we deduce $\top(w) \notin \{+, \star\}$ and hence $\top(w\downarrow) \in \{+, \star\}$. By induction hypothesis applied to w, we would get then $w\downarrow \in G(St_<(w))$ (we can forget the $\overline{F}$ since $\top(w\downarrow) \in \{+, \star\}$). Therefore we can replace w by a (definitely) smaller w_0 to get a smaller u'.

Therefore, by induction hypothesis applied to u', we have $u\downarrow = u'\downarrow \in \overline{F}(G(St_<(u'))\downarrow)$ (*). Now, by hypothesis 1, we have

$$Fact(u') \subseteq \{u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_k\} \cup Fact(w') \cup G(w')$$

Therefore, by hypothesis 4, we get $St_<(u') \subseteq St_<(u) \cup St(w') \cup G(w')$. Let us consider each of the three cases for w' and show that $u\downarrow \in \overline{F}(G(St_<(u))\downarrow)$.

- $w' = w$. In this case, we have $G(St_<(u')) \subseteq G(St_<(u))$ and we conclude from (*).
- $w' = h(w)$. In this case, we have $G(St_<(u')) \subseteq G(St_<(u)) \cup \{w'\}$. By normalizing the two sides, we have $G(St_<(u'))\downarrow \subseteq G(St_<(u))\downarrow \cup \{w'\downarrow\}$. Using $u_i\downarrow = w'\downarrow$ and $u_i \in St_<(u)$, we deduce that $G(St_<(u'))\downarrow \subseteq G(St_<(u))\downarrow$ and conclude by (*).

So we conclude the second case.

We are left with the third case. Suppose first that

$$v = exp(exp(\ldots exp(h(u\downarrow), v_1), \ldots, v_m)$$

and $\top(u\downarrow) = +$. Then $v\downarrow = h(u\downarrow \star v_1 \star \ldots \star v_m\downarrow)\downarrow$. So we obtain $h(u\downarrow \star v_1 \star \ldots \star v_m\downarrow)\downarrow = u\downarrow$. By inspecting the rules of EP, we see that this is not possible: no rule can remove the h apart $h(e_+) \to e_\bullet$ and $\top(e_\bullet) = \bullet$.

Suppose now that $v = h(u\downarrow) \bullet v'$ and $\top(u\downarrow) = \star$. Note that, in this case, $\top(u) = \top(v) = \bullet$. Let j, $1 \leq j < n$, be s.t. $\forall i \leq j.u_i\downarrow = h(u\downarrow)$ and $\forall i > j.u_i\downarrow \neq h(u\downarrow)$. We have then $h(\zeta_+[u\downarrow]) \bullet \zeta'_\bullet[u_{j+1}\downarrow, \ldots, u_k\downarrow]\downarrow = u\downarrow$. Therefore, $\zeta'_\bullet[u_{j+1}\downarrow, \ldots, u_k\downarrow]\downarrow = u\downarrow \bullet h(-\zeta_+[u\downarrow])$.

Since $\top(u\downarrow) = \star$, we get $u\downarrow \in \mathrm{Fact}(u\downarrow \bullet h(-\zeta_+[u\downarrow]))$. Therefore, by hypothesis 2, we get $u\downarrow \in \mathrm{Fact}(\zeta'_\bullet[u_{j+1}\downarrow, \ldots, u_k\downarrow])$. Now, by the definition of j, for all $j+1 \leq i \leq k$, we have $u\downarrow \notin G(u_i\downarrow)$. Therefore, by lemma 28, we get:

- either $\exists i.u\downarrow = u_i\downarrow$;
- or $\exists i.u\downarrow \in \mathrm{Fact}(u_i\downarrow)$ and $\top(u_i\downarrow) \neq \top(u_i)$;

In the first case, we conclude immediately. In the second, we conclude as above. Therefore we conclude the proof of the lemma.

Corollary 4. *For every term u s.t. $\top(u\downarrow) \neq \top(u)$, there is a $w \in G(St_<(u))$ s.t. $u\downarrow \in \{w\downarrow, h(w\downarrow)\}$. Moreover, we have $u \neq h(w)$.*

Proof. By lemma 29, we know that such a w exists. Assume by absurd that $u = h(w)$. Then, since $w \in G(\mathrm{St}_<(u))$, we must have $\top(w) \in \{exp, \bullet, h\}$. By inspecting the rules of $\mathcal{R}_{EP}$, we deduce that $\top(w\downarrow) \in \{exp, \bullet, h\}$. So we deduce $\top(u\downarrow) = \top(u) = h$, contradicting the hypothesis.

Let an inlined solution of the system be a substitution σ such that, for every constraint $T \Vdash_\circ x$, there is a pure recipe $\zeta_\circ$ such that $x\sigma = \zeta_\circ(T)$. Note that we do not normalize the terms here.

Lemma 30. *For every term $u \in St(C)$ and every (inlined) θ, for all $w \in G(St_<(u\theta))$, there is a $v \in \overline{G}(St(C)), v < u$ s.t. $w = v\theta$.*

Proof. Let $C = \{T_1 \Vdash x_1, \ldots, T_n \Vdash x_n\} \cup S$ and $u \in St(T_i, x_i)$.

We proceed the proof by induction on i, $1 \leq i \leq n$. If $i = 1$, and $u \in St(T_i)$, u is ground and the lemma holds immediately. If $u = x_1$, we have $u\theta = \zeta[u_1, \ldots, u_k]$, for some $u_1, \ldots, u_k \in T_1$. By hypothesis 1, we get $Fact(\zeta[u_1, \ldots, u_k]) \subseteq Fact(u_1, \ldots, u_k) \cup G(u_1, \ldots, u_k)$ and therefore $G(St_<(u\theta)) \subseteq G(St(T_1))$, $T_1 < u$ and we conclude.

Suppose now that $i > 1$. Let $u \in St(T_i)$. We proceed by induction on the size of u. If u is a variable, we conclude by induction hypothesis. If u is a constant, $St_<(u\theta) = \emptyset$ and the lemma trivially holds. Otherwise, let $u = \zeta[u_1, \ldots, u_k], Fact(u) = \{u_1, \ldots, u_k\}$. By hypothesis 1, we have $Fact(u\theta) \subseteq Fact(u_1\theta, \ldots, u_k\theta) \cup G(u_1\theta, \ldots, u_k\theta)$ and hence, by hypothesis 4, $St_<(u\theta) \subseteq St(u_1\theta, \ldots, u_k\theta) \cup G(u_1\theta, \ldots, u_k\theta)$ and $G(St_<(u\theta)) \subseteq G(St_<(u_1\theta, \ldots, u_k\theta)) \cup G(u_1\theta, \ldots, u_k\theta)$. By induction hypothesis, we get $\forall j. w \in G(St_<(u_j\theta)) \implies \exists v \in \overline{G}(St(C)), v < u_j < u : w = v\theta$. We are left to consider the case $w \in G(u_j\theta)$, for some j. If $w = u_j\theta$, we take $v = u_j$ and we conclude.

Let $w \in G(u_j\theta) \backslash \{u_j\theta\}$. Suppose first that u_j is not a variable. Then $u_j = h(u'_j)$, $w = u'_j\theta$ and $u'_j \in \overline{G}(St_<(u))$, since $u_j \in St_<(u)$. So we can take $v = u'_j \in \overline{G}(St_<(u))$, since also $u'_j < u_j < u$. Otherwise, if u_j is a variable, let $x_{i'}$, for some $i' < i$, be the smallest variable s.t. $x_{i'}\theta = u_j\theta$. We have $u_j\theta = \zeta'[v_1\theta, \ldots, v_m\theta]$, for some $v_1, \ldots, v_m \in T_{i'}$ and pure ζ'. If ζ' is not a variable, then $top(\zeta') \neq h$ and hence $G(u_j\theta) = \{u_j\theta\}$, contradicting the choice of w. Therefore, ζ' is necessarily a variable and we obtain $w \in G(v_l\theta) \backslash \{v_l\theta\}$, for some l. Since x'_i is minimal, v_l is not a variable. Hence we have $v_l = h(v_{l'})$ and we can take $v = v_{l'} \in \overline{G}(St(C))$ and $v'_l < u$, to conclude.

Now let $u = x_i$. We have then $u\theta = \zeta[u_1\theta, \ldots, u_k\theta]$, for some pure ζ and $u_1, \ldots, u_k \in T_i$. Therefore we can conclude as above, by using the hypotheses 1 and 4.

Lemma 31. *For every term $u \in St(C)$ and every inlined solution θ, if $\top(u\theta\downarrow) \neq \top(u\theta)$, there is a $v \in \mathcal{D}(St(C)), v < u$ s.t. $u\theta\downarrow = v\theta\downarrow$.*

Proof. By lemma 29, we get $u\theta\downarrow \in \overline{F}(w\downarrow) = \{w\downarrow, h(w\downarrow)\}$, for some $w \in G(\mathrm{St}_<(u\theta))$. By lemma 30, there is a $v' \in \overline{G}(\mathrm{St}(C)), v' < u$, s.t. $w = v'\theta$. Since, by corollary 4, $u\theta \neq h(w)$, we deduce that $u \neq h(v')$. Therefore $u > \{v', h(v')\}$ and we can choose $v = v'$, if $u\theta\downarrow = w\downarrow$, or $v = h(v')$, if $u\theta\downarrow = h(w\downarrow)$.

As a consequence, we get:

Lemma 7 *For every $u \in St(C)$ and every solution θ, if $\top(u\theta\downarrow) \neq \top(u)$, there is a $v \in \mathcal{D}(St(C))$, $v < u$ s.t. $u\theta\downarrow = v\theta\downarrow$.*

Proof. Note that, since θ satisfies the top constraints on variables of C, we have $\top(u) = \top(u\theta)$. Let θ' be the inlined solution that corresponds to θ. If $\top(u\theta) = \top(u\theta')$, we conclude by lemma 31, since we have $\top(u\theta') \neq \top(u\theta\downarrow) = \top(u\theta'\downarrow)$. Assume now that $\top(u\theta) \neq \top(u\theta')$. This can happen only if $u = x$ or $u = h(x)$, for some $x \in var(C)$. Then $u\theta = x\theta'\downarrow$ or $u\theta = h(x\theta'\downarrow)$. In the first case, $u\theta$ is in normal form and hence we contradict $\top(u) \neq \top(u\theta\downarrow)$. In the second case, either $u\theta$ is in normal form, and we conclude by contradiction, or else $u\theta = e_\bullet$. For handling this last case, we may assume that constants from $\mathcal{F}_{EP}$ are in $\mathrm{St}(C)$.

F.4 Stabilizing the root symbol: replacements in the constraints

The goal of the first transformation step of this part is to reduce the constraint systems to constraint systems in which solutions σ are such that, for any subterm of the constraint, $\top(u\sigma\downarrow) = \top(u)$: the head symbol is stable.

So, in case the top symbol of u is not stable by substitution and rewriting, we know that $u\sigma\downarrow$ must be equal to some term $v \in \mathcal{D}(\mathrm{St}(C))$. Thanks to the equality guessing part, $u =_{\mathcal{S}} v$ if $\mathcal{S}$ is the equational part of C. Now, choosing an appropriate well-founded ordering on terms, we can ensure $u > v$ and replace u with v. Iterating this procedure yields a system in which all subterms have a stable top symbol.

For each maximal $u \in \mathrm{St}(C)$ such that there is a $v \in \mathcal{D}(\mathrm{St}(C))$ such that $u =_{\mathcal{S}} v$ and $u > v$, replace u with v.

Lemma 32. *In the above transformation $C \to C'$, we keep the following invariants:*

- $\forall w \geq u$, $w \in St(C')$, $\top(w\sigma\downarrow) = \top(w\sigma)$
- *if* $u_1, u_2 \in \mathcal{D}(St(C'))$ *and* $u_1\sigma\downarrow = u_2\sigma\downarrow$, *then* $u_1 =_{\mathcal{S}} u_2$.

Proof. Let $w \in \mathrm{St}(C')$ be s.t. $w \geq u$. There are two possibilities: either $w \in \mathrm{St}(C)$ or else $w \in \mathrm{St}(C')\backslash\mathrm{St}(C)$. In the first case, $\top(w\sigma\downarrow) = \top(w\sigma)$ holds by the choice of u.

Assume now, by contradiction, that we are in the second case and $\top(w\sigma\downarrow) \neq \top(w\sigma)$. Then there is a $w' \in \mathrm{St}(C)$ s.t. $w' = \chi[u]$ and $w = \chi[v]$, for some context χ. Note that, since σ is a solution of $\mathcal{S}$ and $u =_{\mathcal{S}} v$, we have by convergence $w'\sigma\downarrow = w\sigma\downarrow$ (and hence $\top(w'\sigma\downarrow) = \top(w\sigma\downarrow)$). Let us show that also $\top(w'\sigma) = \top(w\sigma)$. Indeed, by the definition of $\top$, this is not the case only when $\chi \in \{\epsilon, h(.)\}$. But then, by the definition of the ordering, it can not be the case that $w = \chi[v] \geq u$, when $v < u$. This contradicts the choice of w. So we deduce $\top(w'\sigma) = \top(w\sigma)$ and therefore $\top(w'\sigma\downarrow) \neq \top(w'\sigma)$.

To contradict the maximality of u, it is sufficient now to show that $w' > u$. Indeed, recall that $w' = \chi[u], \chi[v] > u$ and $v < u$, for some $\chi \notin \{\epsilon, h(.)\}$. Therefore, by the definition of the ordering, we have $w' > u$ and thus we conclude the proof of the first part of the lemma.

Let us prove the second part now. If $u_1, u_2 \in St(C')$, the property follows by hypothesis. Otherwise, we have $u_1 = v_1[v]$ and $u_2 = v_2[v]$, with $v_1[u], v_2[u] \in St(C)$. Therefore, $v_1[u] =_S v_2[u]$ and since $u =_S v$, we conclude $u_1 = v_1[v] =_S v_2[v] = u_2$.

As a consequence, we may now assume that all solutions preserve the top symbols of every subterm of the (decuction part of the) constraint:

Lemma 8. *The above transformation $C \to C'$ preserves the solutions and, for every solution σ of C', for every $u \in St(C')$, $\top(u\sigma\downarrow) = \top(u)$.*

Example 17. Let:

$$C = \begin{cases} h((a+b)\star c), c & \Vdash_+ x \\ h((a+b)\star c), c, exp(x-c, c^{-1}) & \Vdash_\bullet y \end{cases}$$

Suppose that $x\sigma = h((a+b)\star c)+c$. Then $exp(x-c,c^{-1})\sigma\downarrow = h(a+b)$ and hence a and b are factors of $y\sigma$. We compute this by guessing $exp(x-c,c^{-1}) = h(a+b)$ and replacing $exp(x-c,c^{-1})$ by $h(a+b)$.

F.5 Eliminating variables from the left hand sides

Lemma 9 *The transformations of section 6.4 are correct and complete and in the resulting constraint system C':*

- *if $T_i \Vdash_\circ x_i \in C'$ and $\circ \in \{+,\star\}$, then for every $x \in Fact_\circ(T_i)$, $H(x) \neq \circ$.*
- *if $[T_i', T_i''] \Vdash_\bullet x_i \in C'$, then:*
  - *for every $x \in Fact_+(T_i'')$, $H(x) \neq +$.*
  - *for every $u = u_1 \bullet \ldots \bullet u_n \bullet h(v_1 + \ldots + v_m) \in T_i'$, $Fact_\bullet(u) = \{u_1,\ldots,u_n,v_1,\ldots,v_m\}$, for every $x \in \mathcal{X}$, $x \in \{u_1,\ldots,u_n\} \implies H(x) \neq \bullet$ and $x \in \{v_1,\ldots,v_m\} \implies H(x) \neq +$.*

Proof. **A.** Let us first show the properties of the system. Let $\circ \in \{+,\star\}$, $T_i \Vdash_\circ x_i \in C'$ and $x \in \mathrm{Fact}_\circ(T_i)$. We show that $H(x) \neq \circ$.

By contradiction, assume that $H(x) = \circ$ and let T_x be a minimal left hand side such that $T_x \Vdash_{\circ'} x \in C'$. By construction of C', $\circ \neq \circ'$. Now, if θ is an inlined solution of C', $x\theta = \zeta_{\circ'}[T_x\theta]$ and $\circ = H(x) = \top(x\theta\downarrow) \neq \top(x\theta) = \circ'$, which contradicts $\forall u \in St(C').\top(u\theta\downarrow) = \top(u\theta)$, an invariant of the previous transformation step.

The second case, when $\circ = \bullet$, follows in the same way.

B. Now we show correctness and completeness: a substitution σ is a solution of a pure constraint system C if and only if σ is a solution of $\overline{C}$.

It is sufficient to show that, for all i, $T_i\sigma \vdash_\circ \overline{T_i}\sigma$ and $\overline{T_i}\sigma \vdash_\circ T_i\sigma$.

We do this by induction on i. When $i = 0$, $\overline{T_i} = T_i$, so we conclude immediately.

For the induction step, suppose first $\circ \in \{+,\star\}$. Let $u = \zeta[u_1,\ldots,u_k,x_{i_1},\ldots,x_{i_p}] \in T_i$ and $\overline{u} = \zeta[u_1,\ldots,u_k,e_\circ,\ldots,e_\circ]\downarrow \in \overline{T_i}$. By origination, the definition of $\overline{u}$ and induction hypothesis, we have $T_i\sigma \vdash_\circ x_{i_j}\sigma$ and $\overline{T_i}\sigma \vdash_\circ x_{i_j}\sigma$, for all $1 \leq j \leq p$. Therefore, it suffices to show that $\overline{u}, x_1, \ldots, x_p \vdash_\circ u$ and $u, x_1, \ldots, x_p \vdash_\circ \overline{u}$. This follows easily by the definition of $\overline{u}$: one just adds back (to prove u from $\overline{u}$), or retrieves (to prove $\overline{u}$ from u) some appropriate term $\zeta_\circ[x_1,\ldots,x_p]$, relying on *AG* properties of $\circ$.

In the other case, when $\circ = \bullet$ and $[T_i,T_i] \Vdash_\bullet x_i \in \mathcal{C}$, let $\overline{[T_i,T_i]} = [T_i',T_i'']$. By the construction of T_i' and T_i'' we can show similarily as above that $T_i\sigma \vdash_\bullet T_i'\sigma; T_i'\sigma \vdash_\bullet T_i\sigma$ and $T_i\sigma \vdash_+ T_i''\sigma; T_i''\sigma \vdash_+ T_i\sigma$. We detail this only for T_i': let $u = \zeta_\bullet[u_1,\ldots,u_k,x_1,\ldots,x_p] \bullet h(\zeta_+[v_1,\ldots,v_m,y_1,\ldots,y_l])$ be s.t.

$$\overline{u} = \zeta'_\bullet[u_1,\ldots,u_k] \bullet h(\zeta'_+[v_1,\ldots,v_m]).$$

By definition of $\overline{u}$ and origination, it is sufficient to show that

$$\zeta'_\bullet[u_1,\ldots,u_k], x_1, \ldots, x_p \vdash_\bullet \zeta_\bullet[u_1,\ldots,u_k,x_1,\ldots,x_p]$$
$$\zeta_\bullet[u_1,\ldots,u_k,x_1,\ldots,x_p], x_1, \ldots, x_p \vdash_\bullet \zeta'_\bullet[u_1,\ldots,u_k]$$

and

$$\zeta'_+[v_1,\ldots,v_m], y_1,\ldots,y_l \vdash_+ \zeta_+[v_1,\ldots,v_m,y_1,\ldots,y_l]$$
$$\zeta_+[v_1,\ldots,v_m,y_1,\ldots,y_l], y_1,\ldots,y_l \vdash_+ \zeta'_+[v_1,\ldots,v_m]$$

This is follows relying on the properties of AG.

It follows that $\overline{[T_i,T_i]}\sigma \vdash x_i\sigma$ iff $[T_i,T_i]\sigma \vdash x_i\sigma$.

Example 18. Let:

$$C = \begin{cases} a & \Vdash_+ x \\ a, b + x & \Vdash_+ y \end{cases}$$

This system has the following equivalent formulation, in terms of diophantine equations:

$$\begin{cases} x = \lambda a \\ y = \lambda' a + \lambda'' b + \lambda'' \lambda' a \end{cases}$$

This system is diophantine. In order to overcome this, we use the monotonicity and origination properties to turn the constraint system into an equivalent one, whose corresponding diophantine system is linear:

$$C' = \begin{cases} a & \Vdash_+ x \\ a, b & \Vdash_+ y \end{cases}$$

Follows an example that illustrates the particular treatment of $\bullet$ in this step. Let

$$C = \begin{cases} a & \Vdash_+ x \\ a & \Vdash_\bullet y \\ a, x + b, h(x + b), y \bullet b & \Vdash_\bullet z \end{cases}$$

The corresponding system of equations is:

$$\begin{cases} x = \lambda a \\ y = \beta a \\ z = z_1 \bullet h(z_2) \\ z_1 = \gamma_1 a \bullet \gamma_2(x + b) \bullet \gamma_4 y \bullet \gamma_3 b \\ z_2 = \theta_1 a + \theta_2 x + \theta_2 b + \theta_3 h(x + b) + \theta_4(y \bullet b) + \gamma_3 x + \gamma_3 b \end{cases}$$

One can see that this system is not linear due to the terms: $\gamma_4 y$, $\theta_2 x$ and $\gamma_3 x$. At the same time, the terms where these variables occur have occurrences that are alien to the theory of some equation, and thus not being subject to decomposition by erasure of variables. This leads us to consider two copies of each term: one for each theory, $\bullet$ and $+$. The system becomes:

$$C = \begin{cases} a & \Vdash_+ x \\ a & \Vdash_\bullet y \\ [a, x + b, h(x + b), y \bullet b; a, x + b, h(x + b), y \bullet b] & \Vdash_\bullet z \end{cases}$$

Now we can erase from each copy the occurrences of variables that give raise to non-linearity. As expected, 3 occurrences have been erased:

$$C = \begin{cases} a & \Vdash_+ x \\ a & \Vdash_\bullet y \\ [a, x + b, h(b), b; a, b, h(x + b), y \bullet b] & \Vdash_\bullet z \end{cases}$$

The corresponding system is now linear.

F.6 Solving equations

Step 1: apply the finite variant property. We can compute a finite set of substitutions $\theta_1, \ldots, \theta_k$ s.t.:
- For every solution σ of E modulo EP, there is an $1 \leq i \leq k$ and a substitution θ s.t. $\sigma = \sigma_i\theta$ and θ is a solution of $E\sigma_i\downarrow$, modulo $AC(+, \star, \bullet)$.
- For every solution, modulo $AC(+, \star, \bullet)$, θ of $E\sigma_i\downarrow$, $\sigma_i\theta$ is a solution of E modulo EP.

In the following, we call by E one of the $E\theta_i$'s.

Step 2: guess the equalities. Now, since new equalities can be introduced in step 1, we guess once more the equalities between the subterms of E.

Step 3: choose ordering and theory indices. We choose a strict linear ordering on the variables of the system and, for each term, we choose a theory in $\{+, \star, \bullet\}$.

Step 4: treat the diophantine equations. We transform the simple equations into pure equations, following the combination algorithm. Now we show how do we treat the formal equations, to reduce them to pure diophantine equations.

Our system E_1 of formal equations now contains equations of the following form:

$$\beta_1 u_1 \circ \ldots \circ \beta_k u_k = \lambda_1 \alpha_1 t_1 \circ \ldots \circ \lambda_n \alpha_n t_n, \forall i. H(t_i) \neq \circ$$

Note that this equation is satisfied only if, for each $1 \leq j \leq k$, u_j is either equal to a t_i, or else u_j is a variable. Indeed, if u_j is not a variable, since (by step 1) $top(u_j\sigma\downarrow) = top(u_j) \neq \circ$ and $\forall i. top(t_i\sigma\downarrow) \neq \circ$, there must be a t_i s.t. $u_j\sigma\downarrow = t_i\sigma\downarrow$. Since we have guessed equalities, this can happen only when $t_i = u_j$. For all $1 \leq i \leq n$, we let γ_i bet the number of u_j's that are equal to t_i.

Further, letting $\{u_1, \ldots, u_k\} = \{z_1, \ldots, z_m, u_{m+1}, \ldots, u_k\}$ s.t. $\{u_1, \ldots, u_k\} \cap \mathcal{X} = \{z_1, \ldots, z_m\}$, we write

$$\begin{cases} z_1 = \mu_{1,1} t_1 \circ \ldots \circ \mu_{n,1} t_n \\ \ldots \\ z_m = \mu_{1,m} t_1 \circ \ldots \circ \mu_{n,m} t_n \end{cases}$$

Therefore our system is equivalent to the following Diophantine system:

$$\left\{ \alpha_1\lambda_1 = \beta_1\mu_1^1 + \ldots + \beta_m\mu_1^m + \gamma_1 \ldots \alpha_n\lambda_n = \beta_1\mu_n^1 + \ldots + \beta_m\mu_n^m + \gamma_n \right.$$

Applying this to all the formal equations, we get a conjunction of linear diophantine systems, which we can solve by standard methods.

G Proof of locality for *EP*.

Recall that $\overline{F}(t) = \{t, h(t)\}$. We prove in the following the F'-locality of EP, for $F'(T) = \overline{F}(G(T))$ and $\mathrm{St} = \mathrm{St}_{EP}$. Let us give some examples first.

Example 19 (Locality).

1. Let $\sigma = \{x_1 \mapsto h(h(a) - b), x_2 \mapsto b\}$ and let $t = h(h(a) \star b)$. The minimal recipe ζ s.t. $\zeta\sigma\downarrow = v$ is $\zeta = exp(x_1 \bullet h(x_2), x_2)$. If we consider $\zeta' = x_1 \bullet h(x_2) \in \mathrm{St}(\zeta)$, we get $\zeta'\sigma\downarrow = h(h(a))$, which is not in $\mathrm{St}(\sigma, t)$. However, $\zeta'\sigma\downarrow \in \overline{F}(\mathrm{St}(\sigma, t))$.
2. Let $\sigma = \{x_1 \mapsto h(a+b), x_2 \mapsto h(-b), x_3 \mapsto c, x_4 \mapsto h(J_\star(c)\}$, $\zeta = exp(x_4, (x_1 \bullet x_2) \star x_3) \bullet x_2$ and $t = \zeta\sigma\downarrow = h(h(a) - b)$. On one hand, ζ is local only with respect to $\overline{F}(\mathrm{St}(\sigma, t))$, but not with respect to $\mathrm{St}(\sigma, t)$. On the other hand, even the minimal recipe $h(x_1 \bullet x_2) \bullet x_2$ for t, is local only w.r.t $\overline{F}(\mathrm{St}(\sigma, t))$:

$$\dfrac{\dfrac{h(J_\star(c))\quad \dfrac{\dfrac{h(a+b)\quad h(-b)}{h(a)}\bullet \quad c}{h(a)\star c}\star}{h(h(a))}exp \qquad h(-b)}{h(h(a)-b)}\bullet \quad\Longrightarrow\quad \dfrac{\dfrac{\dfrac{h(a+b)\quad h(-b)}{h(a)}\bullet}{h(h(a))}h \qquad h(-b)}{h(h(a)-b)}\bullet$$

3. Above we had a special case. We show here that non-minimal recipes are usually not local: the recipe on the left needs to be replaced with the minimal recipe on the right to get a local proof.

$$\dfrac{h(a)\quad \dfrac{\dfrac{a+b+c\quad -a}{b+c}+\quad J_\star(a)}{(b+c)\star J_\star(a)}\star}{h(b+c)}exp \qquad h(d)}{h(b+c+d)}\bullet \quad\Longrightarrow\quad \dfrac{\dfrac{\dfrac{a+b+c\quad -a}{b+c}+}{h(b+c)}h \qquad h(d)}{h(b+c+d)}\bullet$$

In the first case, $b + c$ is among the subterms of the proof; in the second, it is not: the proof is indeed less complex.

Note that we do not have examples of the necessity of G. It is necessary however in our proof, as we will reuse results from previous sections.

Definition 8. *Let $\zeta[x_1, \ldots, x_k]$ be a pure linear recipe and σ be a substitution. By lemma 28, we have $Fact(\zeta\sigma) \subseteq F_1 \cup \ldots \cup F_k$, where for all i, either $F_i = Fact(x_i\sigma)$ or else $F_i = \{t\} \in G(x_i\sigma)$. If we are in the first case, we call the pair $(x_i, \{x_i \mapsto x_i\sigma\})$* theory preserving. *If we are in the second case, we call it* alien. *For an alien pair $(x_i, \{x_i \mapsto x_i\sigma\})$ we call the term $t \in G(x_i\sigma) \cap Fact(\zeta\sigma)$ an alien* witness.

Lemma 33. *If ζ is a recipe with $Fact(\zeta) = \{x_1, \ldots, x_k\} \subseteq \mathcal{X}$ and $t, v_1, \ldots, v_k$ are terms, then $\zeta\{x_1 \mapsto v_1, \ldots, x_k \mapsto v_k\}[t \mapsto c] = \zeta[v'_1, \ldots, v'_k]$ where*

- *$v'_i = v_i[t \mapsto c]$ if $v_i \notin F(t)$*
- *$v'_i \in \{v_i, v_i[t \mapsto c], \chi[c]\}$ if $v_i = \chi[t] \in F(t)$. Moreover, if t is an alien witness for $(x_i, \{x_i \mapsto v_i\})$, then $v'_i = \chi[c]$.*

Proof. From (**1**),

$$Fact(\zeta\{x_1 \mapsto v_1, \ldots, x_k \mapsto v_k\} \subseteq \{v_1, \ldots, v_k\} \cup Fact(v_1, \ldots, v_k) \cup G(v_1, \ldots, v_k)$$

Moreover, $t \in St(G(v_i))$ iff $t \in St(v_i)$ or $v_i \in F(t)$ since, by hypothesis **4**, for any v, $Fact(F(v)) = Fact(v)$ hence for any u, $Fact(G(u)) = Fact(u)$. It follows that, for every i, either $v_i \notin F(t)$, in which case $\zeta\{x_i \mapsto v_i\}[t \mapsto c] = \zeta\{x_i \mapsto v_i[t \mapsto c]\}$, or else $v_i \in F(t)$.

In the latter case, suppose first that $(x_i, \{x_i \mapsto v_i\})$ is theory preserving. That is, $Fact(\zeta\{x_i \mapsto v_i\}) = Fact(v_i) \cup \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k\}$. Hence we have $\zeta\{x_i \mapsto v_i\}[t \mapsto c] = \zeta\{x_i \mapsto v_i\}$, since $Fact(v_i) = Fact(F(t)) = Fact(t)$, by hypothesis 4, and thus $t \notin St(Fact(v_i))$.

Suppose now that $(x_i, \{x_i \mapsto v_i\})$ is an alien pair. If t is it's alien witness, we have $Fact(\zeta\{x_i \mapsto v_i\}) = \{t\} \cup \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k\}$. Hence we have $\zeta\{x_i \mapsto v_i\}[t \mapsto c] = \zeta\{x_i \mapsto \chi[c]\}$, where $v_i = \chi[t]$.

We are left when t is not the alien witness: for some $t' \neq t$, $t' \in G(v_i)$, we have $Fact(\zeta\{x_i \mapsto v_i\}) = \{t'\} \cup \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k\}$. But then, since, by hypothesis 4, $Fact(v_i) = Fact(t')$, we get $v_i' = v_i[t \mapsto c]$, if $v_i \neq t$, and $v_i' = v_i$, otherwise. So we conclude that $v_i' \in \{v_i, v_i[t \mapsto c], \chi[c]\}$.

Lemma 34. *Let $\zeta(x_1, \ldots, x_n)$ be a pure and linear recipe, σ be a substitution and i an index s.t. $(x_i, \{x_i \mapsto x_i\sigma\})$ is an alien pair, whose alien witness is $x_i\sigma$ (i.e. $x_i\sigma \in Fact(\zeta\sigma)$). Suppose that $\top(x_i\sigma\downarrow) = \top(x_i\sigma)$. Then $(x_i, \{x_i \mapsto x_i\sigma\downarrow\})$ is also an alien pair.*

Proof. This follows by the last part of lemma 28.

Lemma 35. *For every term u and every substitution σ, $St(u\sigma) \subseteq St(u)\sigma \cup St(\sigma) \cup G(\sigma)$.*

Proof. If u is a variable, we conclude. Otherwise, by definition, $\mathrm{St}(u\sigma) = \{u\sigma\} \cup \mathrm{St}(\mathrm{Fact}(u\sigma))$. By hypothesis 1,

$$\mathrm{St}(u\sigma) \subseteq \{u\sigma\} \cup \mathrm{St}(\mathrm{Fact}(u)\sigma) \cup \mathrm{St}(\sigma) \cup \mathrm{St}(G(\sigma))$$

Now, for every variable x, $t \in G(x\sigma)$ iff $x\sigma \in F(t)$ and, from hypothesis 4, $\mathrm{Fact}(x\sigma) = \mathrm{Fact}(t)$. It follows that $\mathrm{Fact}(\sigma) = \mathrm{Fact}(G(\sigma))$, Hence

$$\mathrm{St}(G(\sigma)) = G(\sigma) \cup \mathrm{St}(\mathrm{Fact}(G(\sigma))) = G(\sigma) \cup \mathrm{St}(\mathrm{Fact}(\sigma)) \subseteq G(\sigma) \cup \mathrm{St}(\sigma)$$

Moreover, by induction hypothesis,

$$\mathrm{St}(\mathrm{Fact}(u)\sigma) \subseteq \mathrm{St}(\mathrm{Fact}(u))\sigma \cup G(\sigma) \cup \mathrm{St}(\sigma)$$

We therefore conclude $\mathrm{St}(u\sigma) \subseteq \mathrm{St}(u)\sigma \cup \mathrm{St}(\sigma) \cup G(\sigma)$..

Lemma 36 (Decomposition). *Let σ be a substitution in normal form and ζ be a recipe. Let $v = \zeta\sigma$ and suppose that $\top(v\downarrow) \neq \top(v)$. Then either there is a ζ', $|\zeta'| < |\zeta|$, s.t. $v\downarrow = \zeta'\sigma\downarrow$ or else $v\downarrow \in \overline{F}(G(St(\sigma)))$.*

Example 20. Let us show when we need to apply $\overline{F}$. Take $\sigma = \{x_1 \mapsto h(a + b), x_2 \mapsto h(J_+(b)), x_3 \mapsto c, x_4 \mapsto h(J_\star(c)\}$ and $\zeta = exp(x_4, (x_1 \bullet x_2) \star x_3)$. We have $\top(v) = exp$, $\top(v\downarrow) = h$ and $v\downarrow = h(h(a)) \in \overline{F}(St(\sigma)))$.

Proof. We proceed by induction on the size of ζ. If ζ is a variable, we conclude trivially.

Otherwise, by lemma 29, we have $v\downarrow \in \overline{F}(G(St_<(\zeta\sigma))\downarrow)$.

By lemma 35, we have $St_<(\zeta\sigma) \subseteq St_<(\zeta)\sigma \cup St(\sigma) \cup G(\sigma)$ and therefore, by idempotence and monotonicity of G, $G(St_<(\zeta\sigma)) \subseteq G(St_<(\zeta))\sigma \cup G(St(\sigma))$. Therefore, either $v\downarrow \in \overline{F}(G(St(\sigma)))$, and we conclude, or else $v\downarrow \in \overline{F}(G(St_<(\zeta))\sigma\downarrow)$. Let $\zeta' \in G(St_<(\zeta))$ be s.t. $v\downarrow = \zeta'\sigma\downarrow$ or $v\downarrow = h(\zeta'\sigma\downarrow)$.

If $v\downarrow = \zeta'\sigma\downarrow$, we conclude, since $|\zeta'| < |\zeta|$. Otherwise, $|h(\zeta')| \geq |\zeta|$ only when $\zeta = h(\zeta')$. Suppose hence that $\zeta = h(\zeta')$, with $\zeta' \in St(\zeta)$ and ζ' not a variable. Therefore $\top(\zeta'\sigma) \notin \{+, \star\}$. Let $v' = \zeta'\sigma$. Since $\top(\zeta\sigma\downarrow) \neq \top(\zeta\sigma)$, by definition of $\top$, we must have $\top(v'\downarrow) \neq \top(v')$. Moreover, we can not have $top(v'\downarrow) = h$: otherwise, since $\top(v') \notin \{+, \star\}$, we would have $\top(v) = \top(v\downarrow)$. Therefore, by the induction hypothesis applied to ζ', we get either $v'\downarrow = \zeta''\sigma\downarrow$, for some ζ'' s.t. $|\zeta''| < |\zeta'| < |\zeta|$, or else $v'\downarrow \in \overline{F}(G(St(\sigma))$. Since $top(v'\downarrow) \neq h$, we deduce $v'\downarrow \in G(St(\sigma))$.

Hence, either $v\downarrow = h(\zeta'')\sigma\downarrow$ with $|h(\zeta'')| < |\zeta|$, or else $v\downarrow \in \overline{F}(G(St(\sigma)))$ and we conclude.

Theorem 4 (Locality). *Let σ be a normalized substitution and v be a term s.t. $\sigma \vdash v$. Then there is a recipe ζ s.t.*

- $\zeta\sigma\downarrow = v$
- $\forall \zeta' \in St(\zeta).\zeta'\sigma\downarrow \in \overline{F}(G(St(\sigma))) \cup F(G(St(v)))$.

Proof. Let ζ be the minimal (in size) recipe s.t. $\zeta\sigma\downarrow = v$. We show that ζ satisfies the required properties, by induction. If ζ is a variable, we conclude immediately: $v \in St(\sigma)$.

Otherwise, let $\zeta = \zeta_0[\zeta_1, \ldots, \zeta_k]$, where $Fact(\zeta) = \{\zeta_1, \ldots, \zeta_k\}$. Let $v_1 = \zeta_1\sigma\downarrow, \ldots, v_k = \zeta_k\sigma\downarrow$. By induction hypothesis, we have $\forall i.\forall \zeta' \in St(\zeta_i) : \zeta'\sigma\downarrow \in \overline{F}(G(St(\sigma))) \cup F(G(St(v_i)))$ (*).

Now let $\zeta' \in St(\zeta)$. By definition of subterms, we have $\zeta' \in \{\zeta\} \cup St(\zeta_1, \ldots, \zeta_k)$. If $\zeta' = \zeta$ we conclude, since then $\zeta'\sigma\downarrow = v \in G(St(v))$. Otherwise, by (*), we have $\zeta'\sigma\downarrow \in \overline{F}(G(St(\sigma))) \cup F(G(St(v_1, \ldots, v_k)))$ (**).

We show now that $\forall i.v_i \in \overline{F}(G(St(\sigma))) \cup F(G(St(v)))$.

First, if $\top(v_i) = \top(\zeta_i\sigma\downarrow) \neq \top(\zeta_i\sigma)$, by lemma 36 and the minimality of ζ, we have $v_i \in \overline{F}(G(St(\sigma)))$ (***).

Now, let $Big = \{v_{i_1}, \ldots, v_{i_n}\} \subseteq \{v_1, \ldots, v_k\}$ be the set of terms s.t. $\forall j : v_{i_j} \notin \overline{F}(G(St(\sigma))) \cup F(G(St(v)))$. Suppose by absurd that this set is not empty. By (***), for all the terms v_i in Big, we have $\top(v_i) = \top(\zeta_i\sigma)$. Therefore, since, for all i, $(x_i, \{x_i \mapsto \zeta_i\sigma\})$ is an alien pair for $\zeta_0[x_1, \ldots, x_k]$, whose alien witness is $\zeta_i\sigma$, we deduce, by lemma 34, that $(x_i, \{x_i \mapsto v_i\})$ is also an alien pair. Hence, for each $v_i \in Big$, there is an alien witness t_i. Let t_m be a maximal alien witness in $\{t_{i_1}, \ldots, t_{i_n}\}$.

For every $1 \leq j \leq k$, let us see what value $v_j[t_m \mapsto c]$ can have, for a public constant c. First, if $v_j \notin Big$, then $t_m \notin St(v_j)$. Indeed, otherwise we would have $t_m \in St(\overline{F}(G(St(\sigma))) \cup F(G(St(v)))) \subseteq \overline{F}(G(St(\sigma))) \cup F(G(St(v)))$. Therefore we get $v_m \in F(t_m) \subseteq F(\overline{F}(G(St(\sigma))) \cup F(G(St(v)))) \subseteq \overline{F}(G(St(\sigma))) \cup F(G(St(v)))$, by using the properties of F. So we get a contradiction to $v_m \in Big$. Therefore, if $v_j \notin Big$, we have $t_m \notin St(v_j)$ and $v_j[t_m \mapsto c] = v_j$.

Now suppose that $v_i \in Big$ and thus $v_i \in F(t_i)$, for some alien term t_i. Then, by hypothesis 4, we have $Fact(v_i) = Fact(t_i)$. Therefore, since t_m is a maximal alien term and $t_m \notin Fact(t_i)$, we get $v_i[t_m \mapsto c] = v_i$, if $v_i \neq t_m$, and $v_i[t_m \mapsto c] = c$, otherwise.

Now, note that $v \notin F(t_m)$. Otherwise we would have $v_m \in F(G(v))$, a contradiction to $v_m \in Big$. Therefore, since $v = \zeta_0[v_1, \ldots, v_k]\downarrow$ and $\zeta_0[v_1, \ldots, v_k]$ is with all its factors in normal form, by hypothesis 3, we have $v[t_m \mapsto c] = \zeta_0[v_1, \ldots, v_k][t \mapsto c]\downarrow$. Moreover, $t_m \notin St(v)$ (otherwise, we would have $v_m \in F(G(St(v)))$), therefore $v[t_m \mapsto c] = v$.

Further, by using the lemma 33, we get: $\zeta_0[v_1, \ldots, v_k][t_m \mapsto c] = \zeta_0[v'_1, \ldots, v'_k]$, where:

- if $v_i \notin F(t_m)$, then $v'_i = v_i[t \mapsto c]$.
- if $v_i \in F(t_m)$ and t_m is not the corresponding alien term, then $v'_i \in \{v_i[t \mapsto c], v_i\}$.
- $v'_i = \chi[c]$, otherwise. In particular, $v'_m = \chi[c] \subseteq \{c, h(c)\}$.

Since we have $v_i[t \mapsto c] \in \{v_i, c\}$, for all i, we conclude that we have a proof ζ' of $T \vdash \zeta_0[v'_1, \ldots, v'_k]\downarrow = \zeta_0[v_1, \ldots, v_k][t_m \mapsto c]\downarrow = v$. Moreover, since at least the proof of v_m, ζ_m, was replaced by a smaller, trivial proof in $\{c, h(c)\}$, we have that ζ' is smaller than ζ (if $h(c)$ is not smaller than ζ_m, then $\zeta_m = h(x)$ and $v_m \in F(\sigma)$: contradiction). A contradiction to the minimality of ζ. Hence Big is empty and therefore $\{v_1, \ldots, v_k\} \subseteq \overline{F}(G(St(\sigma))) \cup F(G(St(v)))$. (***)

Finally, putting together (**) and (***) and using the properties of F and G, we conclude $\forall \zeta' \in St(\zeta).\zeta'\sigma\downarrow \in \overline{F}(G(St(\sigma))) \cup F(G(St(v)))$.