Probabilistic Logic over Equations and Domain Restrictions

Andreia Mordido[12] and Carlos Caleiro[13] †

1 SQIG – Instituto de Telecomunicações
2 LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
3 Dep. Mathematics, IST – Universidade de Lisboa, Portugal

We propose and study a probabilistic logic over an algebraic basis, including equations and domain restrictions. The logic combines aspects from classical logic and equational logic with an exogenous approach to quantitative probabilistic reasoning. We present a sound and weakly complete axiomatization for the logic, parameterized by an equational specification of the algebraic basis coupled with the intended domain restrictions. We show that the satisfiability problem for the logic is decidable, under the assumption that its algebraic basis is given by means of a convergent rewriting system, and, additionally, that the axiomatization of domain restrictions enjoys a suitable subterm property. For this purpose, we provide a polynomial reduction to Satisfiability Modulo Theories. As a consequence, we get that validity in the logic is also decidable. Furthermore, under the assumption that the rewriting system that defines the equational basis underlying the logic is also subterm convergent, we show that the resulting satisfiability problem is NP-complete, and thus the validity problem is coNP-complete. We test the logic with meaningful examples in information security, namely by verifying and estimating the probability of the existence of offline guessing attacks to cryptographic protocols.

1. Introduction

The development of formal methods for the analysis of security protocols is a very active research area. Obviously, 'formal methods' should be read as 'logics', but the situation is more complicated. The problem is usually so intricate that suitable logics have not been developed, and the reasoning is usually carried over in an underspecified metalogic, often incorporating ingredients ranging from equational to probabilistic reasoning, from communication and distribution, to temporal or epistemic reasoning (CKW11).

In this paper we present and study a probabilistic logic aimed at dealing with the kind of reasoning used in the verification of security protocols, namely in the analysis of so-called *offline guessing attacks* (Bau05) in a setting where the usual Dolev-Yao intruder (DY83) is extended with some cryptanalytic power (MC09; CBC13). Typically, an attacker eavesdrops on the network and gets hold of a number of messages exchanged by the parties. These messages are usually generated from random data and cyphered using

† Work done under the scope of R&D Unit 50008, financed by the applicable financial framework (FCT/MEC through national funds and when applicable co-funded by FEDER–PT2020). The first author was supported by FCT under the grant SFRH/BD/77648/2011 and by the Calouste Gulbenkian Foundation under *Programa de Estímulo à Investigação* 2011. The second author also acknowledges the support of EU FP7 Marie Curie PIRSES-GA-2012-318986 project GeTFun: Generalizing Truth-Functionality. Work also supported by LASIGE Research Unit, ref. UID/CEC/00408/2019.

secret keys, but often are known to have strong algebraic relationships between them and to comply with certain domain restrictions that may be crucial to the attacker analysis. If the attacker tries to guess the secret keys (a realistic hypothesis in many scenarios, including human-picked passwords, or protocols involving devices with limited computational power) and takes advantage of this knowledge, he may use these relationships to validate his guesses.

The probabilistic logic over equations and domain restrictions (DEqPrL) is designed as a *global* probabilistic logic built on top of a *local* equational base with domain constraints. These two layers are permeated by a quantification mechanism over possible *outcomes* and a quantitative probability operator. Intuitively, we refer to algebraic terms using *names* whose concrete values are gathered in a set of possible *outcomes*, which in turn is endowed with a probability space. The local layer of the logic allows us to reason about equational constraints and domain restrictions on individual outcomes. At the global layer, we can state and reason about qualitative and quantitative properties of the set of all possible outcomes. Not unexpectedly, the quantification we use can be understood as a $S5$-like modality, which also explains why we do not need to consider nested quantifiers. Arguably in the same lines, we will not consider nested probability operators (Pea87). The logic extends the equation-based classical logic of (MC15) with domain restrictions and probabilities. Our approach bears important similarities with exogenous logics in the sense of (MSS05), and with probabilistic logics as developed, for instance, in (FHM90). We provide a sound and weakly complete deductive system for the logic, given a Horn-clause equational specification of the algebraic base and a finite axiomatization for the domain restrictions. We also show that the satisfiability problem for the logic is decidable, under the assumption that its algebraic basis is given by means of a convergent rewriting system and, additionally, that the axiomatization of domain restrictions enjoys a suitable subterm property. We do this by providing a satisfiability algorithm for DEqPrL by means of a polynomial reduction to the Satisfiability Modulo Theories with respect to the theory of quantifier-free linear arithmetic over the integers and reals (QF_LIRA), whose correctness we prove. As a consequence, the validity problem for the logic is also decidable under the same hypothesis. Under the assumption that the rewriting system that defines the equational basis underlying the logic is also subterm convergent, we also show that the resulting satisfiability problem is in NP, and thus the validity problem is in coNP. DEqPrL is used to verify and estimate the probability of the existence of offline guessing attacks to cryptographic protocols.

The paper is outlined as follows: in Section 2 we recall several useful notions of universal algebra and fix some notation on equational reasoning and domain restrictions; in Section 3 we define our logic, its syntax and semantics, and provide a suitable deductive system, whose soundness and (weak) completeness we prove, assuming that we are given a clausal specification of the algebraic basis and a finite axiomatization for domain restrictions; Section 4 is dedicated to showing, by reduction to QF_LIRA, that satisfiability and validity in our logic are decidable whenever the equational basis is given by means of a convergent rewriting system and the axiomatization for domain restrictions enjoys a suitable property; in Section 5 we explore meaningful examples, including an estimation of the probability of offline guessing attacks to simple security protocols; finally, in Section 6, we assess our contributions and discuss future work. Some details of the proofs of our results are given in an Appendix. More details can be found in (Mor17).

2. Preliminaries

In this section we present the technical setting necessary to develop our logic. We begin by recalling some notions of universal algebra and then focus on the details of the semantic structures underlying our logic.

2.1. *Terms and equations*

Let us consider $\mathrm{F} = \{\mathrm{F}_n\}_{n\in\mathbb{N}}$ a $\mathbb{N}$-indexed family of countable sets F_n of function symbols of arity n. Given a set of generators G, we define the set of terms over G, $T_{\mathrm{F}}(G)$, to be the carrier of the free F-algebra $\mathbb{T}_{\mathrm{F}}(G)$ with generators in G. Throughout the text we drop the subscript F when it is clear from context. The set of subterms of a term $t \in T(G)$ is defined as usual and will be denoted by $\mathsf{subtrm}(t)$. Given sets G_1, G_2, a substitution is a function $\sigma : G_1 \to T(G_2)$ that can be easily extended to the set of terms over G_1, $\sigma : T(G_1) \to T(G_2)$.

Fix a countable set of variables X and dub *algebraic terms* the elements of $T(X)$. We denote the set of variables occurring in $t \in T(X)$ by $\mathit{vars}(t)$. Given a F-algebra $\mathbb{A}$ with carrier set A, an assignment is a function $\pi : X \to A$, that is extended as usual to the set of algebraic terms, $[\![\cdot]\!]^{\pi}_{\mathbb{A}} : T(X) \to A$. The set of all assignments is denoted by A^X.

We use $t_1 \approx t_2$ to represent an equation between terms $t_1, t_2 \in T(G)$. The set of all equations over G is denoted by $\mathsf{Eq}(G)$. A Horn clause over G is an expression of the form $(t_1 \approx t'_1\,, \ldots\,, t_k \approx t'_k \Rightarrow t \approx t')$, with $k \geq 0$ and t_1,..., t_k,t'_1,..., $t'_k \in T(G)$. A Horn clause is simply an equation when $k = 0$. We omit the enclosing parentheses when no ambiguities arise. The interpretation of a Horn clause in an algebra $\mathbb{A}$ with respect to $\pi \in A^X$ is defined by: $\mathbb{A}, \pi \Vdash (t_1 \approx t'_1, \ldots, t_k \approx t'_k \Rightarrow t \approx t')$ if whenever $[\![t_i]\!]^{\pi}_{\mathbb{A}} = [\![t'_i]\!]^{\pi}_{\mathbb{A}}$ for each $1 \leq i \leq k$ then $[\![t]\!]^{\pi}_{\mathbb{A}} = [\![t']\!]^{\pi}_{\mathbb{A}}$. An algebra $\mathbb{A}$ satisfies a Horn clause if it is satisfied by $\mathbb{A}$ along with each $\pi \in A^X$. More generally, a Horn clause is satisfied in a class of algebras $\mathcal{A}$ if it is satisfied in every $\mathbb{A} \in \mathcal{A}$. Given a finite set of Horn clauses Γ, the clausal theory of Γ, $\mathsf{Th}(\Gamma)$, is the least set of clauses containing Γ that is stable under reflexivity, symmetry, transitivity and congruence and under application of substitutions. An equational theory is simply a clausal theory where Γ is composed by equations.

We are particularly interested in equational theories generated by convergent rewriting systems. A rewriting system R is a finite set of rewrite rules $l \to r$, where $l, r \in T(X)$ and $\mathsf{vars}(r) \subseteq \mathsf{vars}(l)$. Given a rewriting system R and a set of generators G, the rewriting relation $\to_R \subseteq T(G) \times T(G)$ on $T(G)$ is the smallest relation such that:

- if $(l \to r) \in R$ and $\sigma : X \to T(G)$ is a substitution then $l\sigma \to_R r\sigma$
- if $f \in \mathrm{F}_n$, $t_1, \ldots, t_n, t'_i \in T(G)$ and there exists $i \in \{1, \ldots, n\}$ such that $t_i \to_R t'_i$ then $f(t_1, \ldots, t_i, \ldots, t_n) \to_R f(t_1, \ldots, t'_i, \ldots, t_n)$.

We denote by $\to^*_R$ the reflexive and transitive closure of $\to_R$. R is confluent if, given $t \in T(G)$, $t \to^*_R t'$ and $t \to^*_R t''$ implies that there exists $t^* \in T(G)$ such that $t' \to^*_R t^*$ and $t'' \to^*_R t^*$. R is terminating if there exists no infinite rewriting sequence. R is convergent if it is confluent and terminating. If a rewriting system is convergent then any $t \in T(G)$ has a unique normal form (see (BN99)), i.e., there exists a term $t{\downarrow} \in T(G)$ such that $t \to_R^* t{\downarrow}$ and $t{\downarrow}$ is irreducible. The equational theory generated by a convergent rewriting system R is the relation $\approx_R \subseteq T(G)\times T(G)$ such that $t_1 \approx_R t_2$ if and only if $t_1{\downarrow} = t_2{\downarrow}$, also said to be a convergent equational theory, and is known to always be decidable (see (BN99)). An equational theory is said to be *subterm convergent* if each rule of the underlying rewriting system rewrites to a strict subterm.

Example 2.1. The sum (xor) of single bits can be characterized considering a signature $\mathrm{F}^{\mathsf{xor}}$ with three function symbols: $\mathsf{zero} \in \mathrm{F}_0^{\mathsf{xor}}$, $\mathsf{suc} \in \mathrm{F}_1^{\mathsf{xor}}$, $\oplus \in \mathrm{F}_2^{\mathsf{xor}}$, and the equational theory $\mathsf{Th}(\Gamma^{\mathsf{xor}})$ where $\Gamma^{\mathsf{xor}} = \{\mathsf{zero} \oplus x \approx x, \mathsf{suc}(x) \oplus y \approx x \oplus \mathsf{suc}(y), \mathsf{suc}(\mathsf{suc}(x)) \approx x\}$. Obviously, $\mathbb{Z}_2$ with the usual interpretations for zero, successor and sum modulo 2 satisfies Γ^{xor}. Furthermore, it must be clear that the rewriting system obtained by giving to each of the equations a left-to-right orientation is convergent. However, it is not subterm convergent due to the second equation. △

2.2. *Domain restrictions*

Let $\mathcal{D}$ denote a finite set of domain names. We use $t \in D$ (resp., $t \notin D$) to represent the fact that a term $t \in T(G)$ belongs (resp., does not belong) to a domain $D \in \mathcal{D}$. We dub the expression $t \in D$ (resp., $t \notin D$) a *positive* (resp., *negative*) *domain restriction*. Further, we use $\mathsf{DRes}(G)$ to denote the set of all positive domain restrictions over G. A domain clause is an expression of the form $(t_1 \in D_1, \dots, t_{k_1} \in D_{k_1} \Rightarrow t'_1 \circledcirc D'_1, \dots, t'_{k_2} \circledcirc D'_{k_2})$, where the right-hand side is a non-empty sequence of (positive or negative) domain restrictions, i.e., $k_2 > 0$ and $\circledcirc \in \{\in, \notin\}$. When $t'_1 = \dots = t'_{k_2} = t$ and $t_1, \dots, t_{k_1} \in \mathsf{subtrm}(t)$, we say that the domain clause satisfies the *subterm property*. Again, we omit the enclosing parentheses when no ambiguities arise.

We define an algebraic domain interpretation as a pair $(\mathbb{A}, I^{\mathbb{A}})$, where $\mathbb{A}$ is a F-algebra and $I^{\mathbb{A}} : \mathcal{D} \to 2^A$ fixes an interpretation of domain names as subsets of $\mathbb{A}$. Given an assignment $\pi \in A^X$, the interpretation of domain clauses is defined, as expected, by: $(\mathbb{A}, I^{\mathbb{A}}), \pi \Vdash (t_1 \in D_1, \dots, t_{k_1} \in D_{k_1} \Rightarrow t'_1 \circledcirc D'_1, \dots, t'_{k_2} \circledcirc D'_{k_2})$ if whenever $[\![t_i]\!]^\pi_{\mathbb{A}} \in I^{\mathbb{A}}(D_i)$ for each $1 \le i \le k_1$ then $[\![t'_j]\!]^\pi_{\mathbb{A}} \circledcirc I^{\mathbb{A}}(D'_j)$ for some $1 \le j \le k_2$. An algebraic domain interpretation $(\mathbb{A}, I^{\mathbb{A}})$ satisfies a domain clause if it is satisfied by $(\mathbb{A}, I^{\mathbb{A}})$ along with each $\pi \in A^X$. Moreover, a domain statement is satisfied in a class of algebraic domain interpretations $\mathcal{I}$ if it is satisfied by each $(\mathbb{A}, I^{\mathbb{A}}) \in \mathcal{I}$.

Example 2.2. Let us extend Example 2.1 by introducing a couple of domain names, $\mathcal{D}^{\mathsf{xor}} = \{\mathsf{even}, \mathsf{odd}\}$, which are intended to obey some domain clauses:

$$\Lambda^{\mathsf{xor}} = \{\mathsf{zero} \in \mathsf{even}, (x \in \mathsf{even} \Rightarrow \mathsf{suc}(x) \in \mathsf{odd}), (x \in \mathsf{odd} \Rightarrow \mathsf{suc}(x) \in \mathsf{even}), (x \in \mathsf{odd} \Rightarrow x \notin \mathsf{even})\}.$$

Note that each domain clause in Λ^{xor} satisfies the subterm property, as the behavior of terms is conditioned by restrictions on their subterms. △

3. The logic

In this section we introduce the syntax and semantics of our logic. Then, we define a deductive system for the logic, building upon given clausal specifications of the intended class of algebraic domain interpretations. We conclude by showing soundness and completeness of the deductive system.

3.1. *Syntax*

The logic DEqPrL relies on fixing a signature F, a set of variables X, and a finite set $\mathcal{D}$ of domain names. We also introduce a countable set of *names* N, distinct from algebraic variables. We dub elements of $T(N)$ as *nominal terms*, and let $\mathsf{names}(t)$ stand for the set of names that occur in $t \in T(N)$. Whenever $\mathsf{names}(t) = \varnothing$, the nominal term t is said to be a *nameless term*.

The local language of the logic, designed to express equational constraints and domain restrictions, consists of the set Loc of local formulas defined by the following grammar:

$$\mathsf{Loc} ::= \mathsf{Eq}(N) \mid \mathsf{DRes}(N) \mid \neg\mathsf{Loc} \mid \mathsf{Loc} \wedge \mathsf{Loc} \ .$$

Additionally, we want to express global properties of local formulas, either by quantification or by extracting probabilities. For the purpose, we need a term language Term consisting of linear probabilistic terms with rational coefficients defined by the grammar:

$$\mathsf{Term} ::= \mathbb{Q} \cdot \mathsf{Pr}(\mathsf{Loc}) + \cdots + \mathbb{Q} \cdot \mathsf{Pr}(\mathsf{Loc}) \ ,$$

which we use to define the set Prob of probabilistic statements as follows:

$$\mathsf{Prob} ::= \mathsf{Term} \geq \mathbb{Q} \ .$$

Finally, the language of the logic consists of the following set Glob of global formulas:

$$\mathsf{Glob} ::= \forall\mathsf{Loc} \mid \mathsf{Prob} \mid \neg\mathsf{Glob} \mid \mathsf{Glob} \wedge \mathsf{Glob} \ .$$

Both our local and global languages are to be interpreted classically: the former over an equational base with domain restrictions, and the later over local formulas instead of propositional variables. We abbreviate $\neg(t_1 \approx t_2)$ by $t_1 \not\approx t_2$, $\neg(t \in D)$ by $t \notin D$ for $t, t_1, t_2 \in T(N)$, $D \in \mathcal{D}$, and also use the usual abbreviations: $\psi_1 \vee \psi_2$ abbr. $\neg(\neg\psi_1 \wedge \neg\psi_2)$, $\psi_1 \to \psi_2$ abbr. $\neg\psi_1 \vee \psi_2$, $\psi_1 \leftrightarrow \psi_2$ abbr. $(\psi_1 \to \psi_2) \wedge (\psi_2 \to \psi_1)$, where either $\psi_1, \psi_2 \in \mathsf{Loc}$ or $\psi_1, \psi_2 \in \mathsf{Glob}$; linear probabilistic terms have the common abbreviations saying that $q \cdot (q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell))$ abbr. $(q \cdot q_1) \cdot \mathsf{Pr}(\varphi_1) + \cdots + (q \cdot q_\ell) \cdot \mathsf{Pr}(\varphi_\ell)$, $-q \cdot w$ abbr. $(-q) \cdot w$, $w_1 + w_2$ abbr. $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell) + q'_1 \cdot \mathsf{Pr}(\varphi'_1) + \cdots + q'_\ell \cdot \mathsf{Pr}(\varphi'_\ell)$, whenever w_1 is of the form $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell)$ and w_2 is of the form $q'_1 \cdot \mathsf{Pr}(\varphi'_1) + \cdots + q'_\ell \cdot \mathsf{Pr}(\varphi'_\ell)$; probabilistic formulas result from the usual abbreviations $w_1 \geq w_2 + q$ abbr. $w_1 - w_2 \geq q$, $w < q$ abbr. $\neg(w \geq q)$, $w \leq q$ abbr. $-w \geq -q$, $w > q$ abbr. $-w < -q$, $w = q$ abbr. $w \leq q \wedge w \geq q$, $q_1 \leq w \leq q_2$ abbr. $w \geq q_1 \wedge w \leq q_2$, where $\ell \geq 1$, $\varphi_1, \ldots, \varphi_\ell \in \mathsf{Loc}$, $q, q_1, q_2, \ldots, q_\ell \in \mathbb{Q}, w, w_1, w_2 \in \mathsf{Term}$. We introduce $\top$ for *local true* abbreviating $\varphi \vee \neg\varphi$ for some $\varphi \in \mathsf{Loc}$ and *local false* $\bot$ representing $\neg\top$. We abuse notation and denote the *global true*, $\forall\top$, and *global false*, $\forall\bot$, also by $\top$ and $\bot$.

A *literal* is a global formula in $\forall\mathsf{Loc} \cup \neg\forall\mathsf{Loc} \cup \mathsf{Prob} \cup \neg\mathsf{Prob}$. We say that a global formula is in *disjunctive normal form* (DNF) if it is a disjunction of one or more conjunctions of literals; it is in *conjunctive normal form* (CNF) if it is a conjunction of disjunctions of literals. The language of the logic allows us to make qualitative and quantitative assertions over local formulas. The universal quantification of a local formula expresses the validity of the local formula in all possible situations, whereas a probabilistic statement measures the probability of satisfying local formula(s). Boolean combinations are allowed in both local and global layers. For instance, the formula $(\mathsf{Pr}(\varphi) \leq 2 \cdot \mathsf{Pr}(\psi \wedge \neg\varphi)) \wedge (\forall\neg\psi \to \forall\neg\varphi)$ should be read as: the probability of φ does not exceed twice the probability of $\psi \wedge \neg\varphi$ and, either ψ holds in some situation or else φ never holds. Note that, contrarily to the discussion carried out by Eijck and Schwarzentruber in (VES14), $\forall\varphi$ implies but is not intended to be equivalent to $\mathsf{Pr}(\varphi) = 1$.

Example 3.1. Let us go back to Example 2.2. Given a name $n \in N$, we want to be able to show that a statement like: $\mathsf{Pr}(\, n \in \mathsf{even}\,) = \mathsf{Pr}(\, \mathsf{suc}(n) \in \mathsf{odd}\,) \;\wedge\; \forall(\, \mathsf{zero} \not\approx \mathsf{suc}(\mathsf{zero})\,)$ is a theorem of the logic whose algebraic basis is axiomatized by Γ^{xor} and whose domain restrictions are given by Λ^{xor}. △

We extend the notion of subterm to global formulas in a standard way, and abuse notation by denoting $\mathsf{subtrm}(\Psi) = \bigcup_{\psi\in\Psi} \mathsf{subtrm}(\psi)$, for $\Psi \subseteq \mathsf{Glob}$. Similarly, we generalize the notion of names occurring in a term to local and global formulas. The set of subformulas of either a local or a global formula ψ is defined in the usual way and is denoted by $\mathsf{subform}(\psi)$. As usual, $\mathsf{subform}(\Psi) = \bigcup_{\psi\in\Psi} \mathsf{subform}(\psi)$. Given a nominal term $t_0 \in T(N)$, a set of names $\tilde{n} = \{n_1, \dots, n_k\} \subseteq N$ such that $\mathsf{names}(t_0) \subseteq \tilde{n}$ and $\tilde{t} = \{t_1, \dots, t_k\} \subseteq T(N)$, $[t_0]^{\tilde{n}}_{\tilde{t}}$ is the nominal term obtained by replacing each occurrence of n_i by t_i, $i \in \{1, \dots, k\}$, i.e., $[t_0]^{\tilde{n}}_{\tilde{t}} = \sigma(t_0)$ where σ is a substitution such that $\sigma(n_i) = t_i$ for each i. This notion is easily extended to local formulas.

3.2. *Semantics*

Names can be thought of as being associated to values that are not made explicit, and which are possibly sampled according to some probability distribution. We call *outcome* to each possible concrete assignment of values to names. For this purpose, given a F-algebra $\mathbb{A}$ with carrier set A, we define an outcome as a function $\rho : N \to A$. The set of all outcomes is denoted by A^N. The interpretation of terms $[\![\cdot]\!]^{\rho}_{\mathbb{A}} : T_{\mathrm{F}}(N) \to A$ is defined as usual. Given an algebraic domain interpretation $(\mathbb{A}, I^{\mathbb{A}})$, the *satisfaction relation for local formulas*, $\Vdash_{\mathsf{loc}}$, is defined inductively as follows:

- $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} t_1 \approx t_2$ iff $[\![t_1]\!]^{\rho}_{\mathbb{A}} = [\![t_2]\!]^{\rho}_{\mathbb{A}}$,
- $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} t \in D$ iff $[\![t]\!]^{\rho}_{\mathbb{A}} \in I^{\mathbb{A}}(D)$,
- $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} \neg\varphi$ iff $(\mathbb{A}, I^{\mathbb{A}}), \rho \nVdash_{\mathsf{loc}} \varphi$,
- $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} \varphi_1 \wedge \varphi_2$ iff $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} \varphi_1$ and $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} \varphi_2$.

In order to interpret global formulas we need to fix an intended set of possible outcomes for names and to endow it with a probability space, which is instrumental for evaluating probabilistic statements.

Definition 3.1. A F-*structure* is a tuple $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P})$ where $(\mathbb{A}, I^{\mathbb{A}})$ is an algebraic domain interpretation, and $\mathbb{P} = (S, \mathscr{A}, \mu)$ is a probability space composed by:

- a non-empty set $S \subseteq A^N$ of *possible outcomes*,
- a σ-algebra $\mathscr{A}$ containing the sets of outcomes satisfying each local formula,
$$\{S^{\varphi} \mid \varphi \in \mathsf{Loc}\} \subseteq \mathscr{A}, \text{ with } S^{\varphi} = \{\rho \in S \mid (\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} \varphi\},$$
- a probability measure μ over $\mathscr{A}$.

Given a F-structure $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P})$ with $\mathbb{P} = (S, \mathscr{A}, \mu)$, the *satisfaction relation for global formulas*, $\Vdash$, is defined inductively as follows:

- $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \forall\varphi$ iff $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} \varphi$ for every $\rho \in S$,
- $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_l \cdot \mathsf{Pr}(\varphi_l) \geq q$ iff $q_1 \cdot \mu(S^{\varphi_1}) + \cdots + q_l \cdot \mu(S^{\varphi_l}) \geq q$,
- $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \neg\delta$ iff $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \nVdash \delta$,

- $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \delta_1 \wedge \delta_2$ iff $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \delta_1$ and $(\mathbb{A}, I, \mathbb{P}) \Vdash \delta_2$.

As usual, given $\Delta \subseteq \mathsf{Glob}$ we write $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \Delta$ if $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \delta$ for each $\delta \in \Delta$.

Our logic is parameterized by a choice of intended algebraic domain interpretations.

Definition 3.2. Given a class $\mathcal{I}$ of algebraic domain interpretations, the *semantic consequence relation* of our logic, $\vDash_{\mathcal{I}} \subseteq 2^{\mathsf{Glob}} \times \mathsf{Glob}$, is such that $\Delta \vDash_{\mathcal{I}} \delta$ whenever, for every F-structure $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P})$ with $(\mathbb{A}, I^{\mathbb{A}}) \in \mathcal{I}$, if $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \Delta$ then $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \delta$.

Example 3.2. Independence cannot in general be expressed in our logic, as its language only allows for linear combinations of probabilistic terms. This could be achieved, however, without spoiling too much the nice properties of the logic, by considering coefficients taken from real closed fields, not necessarily from $\mathbb{Q}$, in the lines of (FHM90; MSS05). However, it would result in a double exponential complexity (Sho67), which we would like to avoid. Even so, we can highlight some situations where one can characterize, reason about, or at least approximate the probabilistic behavior of independent formulas.

Verification of the independence of events is easily modeled within our logic: given a F-structure $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P})$, $\varphi, \psi \in \mathsf{Loc}$ are independent if we can find $\alpha, \beta \in \mathbb{Q}$ such that $\beta \neq 0$ and $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \mathsf{Pr}(\varphi \wedge \psi) = \alpha \ \wedge \ \mathsf{Pr}(\psi) = \beta \ \wedge \ \mathsf{Pr}(\varphi) = \frac{\alpha}{\beta}$. More importantly, we can draw some conclusions on the estimation of probabilities by knowing about the independence of some formulas. If φ and ψ are independent, we can model the expected probabilistic behavior of both events with a finite set of properties, defined within the logic: for fixed and appropriately chosen $n, m \in \mathbb{N}$, we can introduce $n \cdot m$ conditions

$\mathsf{Ind}^{\varphi,\psi}_{i,j}$**:** $\quad \mathsf{Pr}(\varphi) = \frac{1}{i} \wedge \mathsf{Pr}(\psi) = \frac{1}{j} \to \mathsf{Pr}(\varphi \wedge \psi) = \frac{1}{i} \cdot \frac{1}{j}$, for $i \in \{1, \dots, n\}$, $j \in \{1, \dots, m\}$.

As an application, we analyze the simpler version of one-time pad encryption scheme which consists of encrypting a secret bit by summing to it an uniformly generated key-bit.

Inspired in Examples 2.1 and 2.2, consider the signature $\mathrm{F}^{\mathsf{xor}}$ and denote by $\mathcal{I}_{(\Gamma^{\mathsf{xor}}, \Lambda^{\mathsf{xor}})}$ the class of algebraic domain interpretations satisfying the axiomatizations Γ^{xor} and Λ^{xor}. Consider a bit s, which will be kept secret as result of its encryption with a key-bit k. The described properties on the estimation of probabilities for the conjunction of independent events enable us to semantically infer that, under the hypothesis that k is uniformly generated and that bits s and k are independent, $\mathsf{Hyp} = \{\mathsf{Pr}(k \approx \mathsf{zero}) = \frac{1}{2}$, $\mathsf{Pr}(k \approx \mathsf{suc}(\mathsf{zero})) = \frac{1}{2}$, $\mathsf{Ind}^{s,k}_{2,2}$, $\forall(s \approx \mathsf{zero} \vee s \approx \mathsf{suc}(\mathsf{zero})), \forall(k \approx \mathsf{zero} \vee k \approx \mathsf{suc}(\mathsf{zero}))\}$, $s \oplus k$ has uniform distribution:

$$\mathsf{Hyp} \vDash_{\mathcal{I}_{(\Gamma^{\mathsf{xor}}, \Lambda^{\mathsf{xor}})}} \left(\mathsf{Pr}(s \oplus k \approx \mathsf{zero}) = \frac{1}{2} \wedge \mathsf{Pr}(s \oplus k \approx \mathsf{suc}(\mathsf{zero})) = \frac{1}{2}\right).$$

Note that we could generalize properties $\mathsf{Ind}^{\varphi,\psi}_{i,j}$ estimating the probability for the conjunction of independent event by squeezing its value. For a fixed $n \in \mathbb{N}$, $q_1, \dots, q_n \in \mathbb{Q}$ such that $q_1 < \dots < q_n = 1$, and independent events $\varphi, \psi \in \mathsf{Loc}$,

$\widetilde{\mathsf{Ind}}^{i_1 j_1}_{i_2 j_2}$**:** $(q_{i_1} \leq \mathsf{Pr}(\varphi) \leq q_{i_2} \ \wedge \ q_{j_1} \leq \mathsf{Pr}(\psi) \leq q_{j_2}) \to q_{i_1} \cdot q_{j_1} \leq \mathsf{Pr}(\varphi \wedge \psi) \leq q_{i_2} \cdot q_{j_2}$,

for $i_1, i_2, j_1, j_2 \in \{1, \dots, n\}$, would model the estimation of bounds of the probabilities for the conjunction of independent events given bounds for the individual probabilities. $\triangle$

3.3. *Deductive system*

In order to obtain a sound and complete deductive system for our logic, we require that the class $\mathcal{I}$ of intended interpretations is such that its algebras are axiomatized by a set Γ of Horn clauses and the corresponding interpretations for domain names are axiomatized by a finite set Λ of domain clauses of algebraic terms. We say that Γ and Λ are *compatible* if $\mathcal{I}_{(\Gamma,\Lambda)} = \{(\mathbb{A}, I^{\mathbb{A}}) \mid \mathbb{A} \Vdash \Gamma \text{ and } (\mathbb{A}, I^{\mathbb{A}}) \Vdash \Lambda\} \neq \varnothing$. Whenever Γ, Λ are not compatible, the set of models is empty and the logic becomes trivial. The interesting cases are, obviously, the ones where the equational theory and the set of domain restrictions are compatible.

Eq1 $\forall(t \approx t)$
Eq2 $\forall(t_1 \approx t_2 \to t_2 \approx t_1)$
Eq3 $\forall(t_1 \approx t_2 \ \wedge\ t_2 \approx t_3 \to t_1 \approx t_3)$
Eq4 $\forall(t_1 \approx t_1' \wedge \ldots \wedge t_n \approx t_n' \to f(t_1, \ldots, t_n) \approx f(t_1', \ldots, t_n'))$

EqC1 $\forall((\varphi_1 \to (\varphi_2 \to \varphi_3)) \to ((\varphi_1 \to \varphi_2) \to (\varphi_1 \to \varphi_3)))$
EqC2 $\forall(\varphi_1 \to (\varphi_2 \to \varphi_1))$
EqC3 $\forall((\neg\varphi_1 \to \neg\varphi_2) \to (\varphi_2 \to \varphi_1))$
EqC4 $\forall(\varphi_1 \to ((\varphi_1 \to \varphi_2) \to \varphi_2))$

DEq $\forall((t_1 \approx t_2 \wedge t_1 \in \mathrm{D}) \to t_2 \in \mathrm{D})$

N1 $\forall(\varphi_1 \wedge \varphi_2) \leftrightarrow (\forall\varphi_1 \wedge \forall\varphi_2)$
N2 $\forall\neg\varphi \to \neg\forall\varphi$
N3 $\neg\forall\varphi \to \forall\neg\varphi$ if $\mathsf{names}(\varphi) = \varnothing$
N4 $\forall(\varphi_1 \leftrightarrow \varphi_2) \to (\forall\varphi_1 \leftrightarrow \forall\varphi_2)$

C1 $\delta_1 \to (\delta_2 \to \delta_1)$
C2 $(\delta_1 \to (\delta_2 \to \delta_3)) \to ((\delta_1 \to \delta_2) \to (\delta_1 \to \delta_3))$
C3 $(\neg\delta_1 \to \neg\delta_2) \to (\delta_2 \to \delta_1)$
C4 $\dfrac{\delta_1 \quad \delta_1 \to \delta_2}{\delta_2}$

P1 $\mathsf{Pr}(\varphi) \geq 0$
P2 $\mathsf{Pr}(\varphi_1 \wedge \varphi_2) + \mathsf{Pr}(\varphi_1 \wedge \neg\varphi_2) - \mathsf{Pr}(\varphi_1) = 0$
P3 $\forall(\varphi_1 \to \varphi_2) \to \mathsf{Pr}(\varphi_2) \geq \mathsf{Pr}(\varphi_1)$
P4 $\mathsf{Pr}(\top) = 1$

I1 $w \geq q \vee w \leq q$
I2 $w \geq q_1 \to w > q_2$, if $q_1 > q_2$
I3 $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell) \geq q \leftrightarrow q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell) + 0 \cdot \mathsf{Pr}(\varphi_{\ell+1}) \geq q$
I4 $((q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell) \geq q) \wedge (q_1' \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell' \cdot \mathsf{Pr}(\varphi_\ell) \geq q')) \to ((q_1 + q_1') \cdot \mathsf{Pr}(\varphi_1) + \cdots + (q_\ell + q_\ell') \cdot \mathsf{Pr}(\varphi_\ell) \geq q + q')$
I5 $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell) \geq q \to (q' \cdot q_1) \cdot \mathsf{Pr}(\varphi_1) + \cdots + (q' \cdot q_\ell) \cdot \mathsf{Pr}(\varphi_\ell) \geq (q' \cdot q)$, for any $q' > 0$
I6 $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell) \geq q \leftrightarrow q_{i_1} \cdot \mathsf{Pr}(\varphi_{i_1}) + \cdots + q_{i_\ell} \cdot \mathsf{Pr}(\varphi_{i_\ell}) \geq q$, for any permutation $(i_1 \cdots i_\ell)$ of $(1 \cdots \ell)$

D(Λ) $\forall(\sigma(s_1) \in \mathrm{D}_1 \wedge \ldots \wedge \sigma(s_{k_1}) \in \mathrm{D}_{k_1}) \to (\sigma(s_1') \circledcirc \mathrm{D}_1' \vee \ldots \vee \sigma(s_{k_2}') \circledcirc \mathrm{D}_{k_2}')$

E(Γ) $\forall(\sigma(s_1) \approx \sigma(s_1') \wedge \ldots \wedge \sigma(s_n) \approx \sigma(s_n') \to \sigma(t) \approx \sigma(t'))$

for $t, t_1, t_2, t_3, \ldots, t_n, t_1', \ldots, t_n' \in T(N)$, $\varphi, \varphi_1, \varphi_2, \varphi_3, \ldots, \varphi_\ell, \varphi_{\ell+1} \in \mathsf{Loc}$, $\delta_1, \delta_2, \delta_3 \in \mathsf{Glob}$, $\sigma \in T(N)^X$, $w \in \mathsf{Term}$, $q, q', q_1, q_2, \ldots, q_\ell, q_1', \ldots, q_\ell' \in \mathbb{Q}$, $(s_1 \in \mathrm{D}_1, \ldots, s_{k_1} \in \mathrm{D}_{k_1} \Rightarrow s_1' \circledcirc \mathrm{D}_{k_1}', \ldots, s_{k_2}' \circledcirc \mathrm{D}_{k_2}') \in \Lambda$ and $(s_1 \approx s_1', \ldots, s_n \approx s_n' \Rightarrow s \approx s') \in \Gamma$.

Fig. 1: The deductive system $\mathcal{H}_{(\Gamma,\Lambda)}$.

The deductive system $\mathcal{H}_{(\Gamma,\Lambda)}$ consists of a number of axioms and a single inference rule C4, *modus ponens*, as shown in Figure 1. The system combines the different dimensions of this logic: axioms Eq1-Eq4 incorporate standard equational reasoning, namely reflexivity, symmetry, transitivity and congruence; EqC1-EqC4 and C1-C4 incorporate classical reasoning for the local and global layers (just note that locally, *modus ponens* becomes axiom EqC4); N1-N4 characterize the relationship between the local and global layers across the universal quantifier; DEq represents syntactically the expected relation between equations and domain restrictions; I1-I6 incorporate properties of inequalities between rational numbers; P1-P4 represent the standard properties of probabilities; axioms E(Γ) incorporate the clausal specification Γ, whereas axioms D(Λ) characterize the constraints for domains given by Λ. We define, as usual, a deducibility relation $\vdash^{\mathrm{F}}_{(\Gamma,\Lambda)}$. We drop the superscript F whenever it is clear from context.

Basic arithmetic properties, such as $0 \cdot \mathsf{Pr}(\varphi) = 0$ or $q_1 \cdot \mathsf{Pr}(\varphi) + q_2 \cdot \mathsf{Pr}(\varphi) = (q_1 + q_2) \cdot \mathsf{Pr}(\varphi)$, are deducible in $\mathcal{H}_{(\Gamma,\Lambda)}$, as well as some expected properties of the probabilistic operator, namely $\forall\varphi \to \mathsf{Pr}(\varphi) = 1$ or $\forall(\varphi_1 \leftrightarrow \varphi_2) \to \mathsf{Pr}(\varphi_1) = \mathsf{Pr}(\varphi_2)$. The logic is an extension of classical logic at both the local and global layers. Namely, it is easy to see that the *deduction metatheorem* holds. Moreover, we can write any local or global formula in *disjunctive normal form (DNF)*. The behavior of implication across the universal quantifier can be deduced and takes the form of theorem:

N $\qquad \vdash_{(\Gamma,\Lambda)} \forall(\varphi_1 \to \varphi_2) \to (\forall\varphi_1 \to \forall\varphi_2)$.

Example 3.3. A standard example of an equational theory used in information security for formalizing (part of) the capabilities of a so-called *Dolev-Yao attacker* (see, for instance, (Bau05; AC06; AC05)) consists in taking a signature F^{DY} with $\{\cdot\}_\cdot, \{\cdot\}_\cdot^{-1} \in \mathrm{F}_2^{\mathsf{DY}}$ representing symmetric encryption and decryption of a message with a key, $\{\!|\cdot|\!\}_\cdot, \{\!|\cdot|\!\}_\cdot^{-1} \in \mathrm{F}_2^{\mathsf{DY}}$ representing asymmetric encryption of a message with a public key or decryption with a private key, $\mathsf{pub}(\cdot), \mathsf{prv}(\cdot) \in \mathrm{F}_1^{DY}$ representing public and private keys for a principal, $(\cdot,\cdot) \in \mathrm{F}_2^{\mathsf{DY}}$ representing message pairing, and $\pi_1, \pi_2 \in \mathrm{F}_1^{\mathsf{DY}}$ representing projections. The equational properties of these operations can be axiomatized by the subterm convergent equational theory

$$\Gamma^{\mathsf{DY}} = \{\{\{x_1\}_{x_2}\}_{x_2}^{-1} \approx x_1, \{\!|\{\!|x_1|\!\}_{\mathsf{pub}(x_2)}|\!\}_{\mathsf{prv}(x_2)}^{-1} \approx x_1, \pi_1(x_1,x_2) \approx x_1, \pi_2(x_1,x_2) \approx x_2\}.$$

Considering a suitable set of domain names, for instance we may take

$$\mathcal{D}^{\mathsf{DY}} = \{\mathsf{sym\_key}, \mathsf{pub\_key}, \mathsf{prv\_key}, \mathsf{principals}, \mathsf{plaintxt}, \mathsf{ciphertxt}, \mathsf{conc}\},$$

we can also impose some usual domain restrictions:

$$\begin{aligned}\Lambda^{\mathsf{DY}} = \{ & (k \in \mathsf{sym\_key}, t \in \mathsf{plaintxt} \Rightarrow \{t\}_k \in \mathsf{ciphertxt}), (k \in \mathsf{sym\_key}, t \in \mathsf{ciphertxt} \Rightarrow \{t\}_k^{-1} \in \mathsf{plaintxt}),\\ & (n \in \mathsf{principals} \Rightarrow \mathsf{pub}(n) \in \mathsf{pub\_key}), \ (n \in \mathsf{principals} \Rightarrow \mathsf{prv}(n) \in \mathsf{prv\_key}),\\ & (t \in \mathsf{plaintxt}, k \in \mathsf{pub\_key} \Rightarrow \{\!|t|\!\}_k \in \mathsf{ciphertxt}), (t \in \mathsf{ciphertxt}, k \in \mathsf{prv\_key} \Rightarrow \{\!|t|\!\}_k^{-1} \in \mathsf{plaintxt}),\\ & (t \in \mathsf{plaintxt}, t' \in \mathsf{plaintxt} \Rightarrow (t,t') \in \mathsf{conc}), \ (t \in \mathsf{conc} \Rightarrow t \in \mathsf{plaintxt}),\\ & (t \in \mathsf{conc} \Rightarrow \pi_1(t) \in \mathsf{plaintxt}), \ (t \in \mathsf{conc} \Rightarrow \pi_2(t) \in \mathsf{plaintxt}) \ \} \ .\end{aligned}$$

The first domain restriction, for instance, is intended to mean that the encryption of a plaintext with a symmetric key should always lead to a ciphertext. As a result, we can deduce from our logic (see the proof in the Appendix) a bound for the probability of an attack to the symmetric scheme:

$$\mathsf{Pr}(k \approx k^*) = q \cdot \mathsf{Pr}(k^* \in \mathsf{sym\_key}) \vdash_{(\Gamma^{\mathsf{DY}},\Lambda^{\mathsf{DY}})} \forall(k^* \in \mathsf{sym\_key}) \to \mathsf{Pr}(\{\{m\}_k\}_{k^*}^{-1} \approx m) \geq q,$$

asserting that even assuming that a guess k^* to the secret key k is indeed a symmetric key, guessing its concrete value is not simpler than decrypting a message encrypted with k. We can also deduce conditions to rule out the possibility of an attack, like

$$\forall(k \in \mathsf{sym\_key} \land m \in \mathsf{plaintxt}) \vdash_{(\Gamma^{\mathsf{DY}},\Lambda^{\mathsf{DY}})} \forall(\{\{m\}_k\}_{k^*}^{-1} \notin \mathsf{plaintxt} \to k \not\approx k^*),$$

which states that whenever an attempt to guess the secret key k leads to a message outside the scope of plaintexts, the value of k has certainly not been guessed correctly. △

3.4. *Soundness and completeness*

We now show that $\mathcal{H}_{(\Gamma,\Lambda)}$ is a sound and weakly complete proof system for the logic based on the classe $\mathcal{I}_{(\Gamma,\Lambda)}$ of algebraic domain interpretations. In contrast to (MC15), the introduction of probabilistic terms over the rationals carries the expected cost of losing the strong version of completeness (see, for instance, (FHM90; MSS05)). Clearly, our semantic consequence is not compact as we have that $\{w \leq \frac{1}{n} \mid n \in \mathbb{N}\} \vDash_{(\Gamma,\Lambda)} w \leq 0$, but $\Delta \nvDash_{(\Gamma,\Lambda)} w \leq 0$ for any finite set $\Delta \subset \{w \leq \frac{1}{n} \mid n \in \mathbb{N}\}$, which implies that our finitary deductive system $\mathcal{H}_{(\Gamma,\Lambda)}$ cannot aim at strong completeness.

Theorem 3.1. $\mathcal{H}_{(\Gamma,\Lambda)}$ is sound, that is, if $\Delta \vdash_{(\Gamma,\Lambda)} \delta$ then $\Delta \vDash_{(\Gamma,\Lambda)} \delta$.

We omit the proof, as it is straightforward to check soundness of each axiom and inference rule against our semantics. The proof of completeness can be found in the Appendix.

Theorem 3.2. $\mathcal{H}_{(\Gamma,\Lambda)}$ is weakly complete, that is, if $\vDash_{(\Gamma,\Lambda)} \delta$ then $\vdash_{(\Gamma,\Lambda)} \delta$.

Proof. As usual, the proof of completeness follows by contrareciprocal and consists in finding a model for the negation of a formula. The proof combines several known techniques, namely in the context of equational logics, first order logic and probabilistic logics. All these components interact in a non-trivial way and should be taken carefully.

The construction of the F-structure starts with the completion of the set of formulas to satisfy with witnesses for negated global formulas through a well-known Henkin construction. We slightly change this construction and take this opportunity to introduce nameless terms that should represent all the non-negative global formulas to satisfy. Once taken its maximal consistent extension, a quotient is made in the set of nameless terms over the extended signature, that collapses congruence classes of terms that should be equal in all possible outcomes. A domain interpretation is then taken accordingly to the aforementioned maximal consistent set. Afterwards, each negated global formula in the maximal consistent set leads to an outcome, assigning to each name the equivalence class of the appropriate constant. The set of all outcomes should not be empty due to the conjugation of the reflexivity axiom Eq1 with axiom N2.

A probability space is then defined in the lines of (Fagin et. al 1990) choosing carefully a set of atoms of interest and using the result of soundness and completeness for the axioms of inequality to define a probability distribution over those atoms.

The verification that we have effectively found a model for the initial formula follows easily from the construction. Details of the proof can be found in the Appendix. □

4. Decidability and Complexity

In general, our logic cannot be expected to be decidable, as equational theories can easily be undecidable (BN99). We show, however, that our logic is decidable if we require the base equational theory to be convergent, and additionally the underlying domain clauses to have the subterm property. With this purpose, our setup is, from now on, that Γ is a convergent equational theory and Λ is a set of domain clauses with the subterm property.

4.1. *Satisfiability*

We devote this subsection to the analysis of the satisfiability problem for DEqPrL (SAT-DEqPrL). The SAT-DEqPrL problem consists in deciding the existence of a model for a global formula. As we often observe in classical propositional logic, we start by analyzing the satisfiability problem for DEqPrL in which the input formula is required to be in CNF, we call it CNFSAT-DEqPrL problem. We provide a reduction of CNFSAT-DEqPrL to Satisfiability Modulo Theories (SMT) (NOT06) and end up using a Tseitin-like transformation to analyse SAT-DEqPrL.

Moving to the propositional context: To describe an algorithm that reduces SAT-DEqPrL to SMT, we translate local formulas to the propositional context. For that, let us consider a set of propositional symbols corresponding to equations between nominal terms $\mathsf{Eq}(N)^{\mathsf{p}} = \{\mathsf{p}_{t_1 \approx t_2} \mid t_1, t_2 \in T(N)\}$ and a set of propositional symbols for domain restrictions $\mathsf{DRes}(N)^{\mathsf{p}} = \{\mathsf{p}_{t \in \mathrm{D}} \mid t \in T(N), \mathrm{D} \in \mathcal{D}\}$, and then define the translation of an arbitrary local formula $\varphi \in \mathsf{Loc}$ to a propositional formula prop_φ inductively, by:

- if φ is of the form $t_1 \approx t_2$, prop_φ is precisely $\mathsf{p}_{t_1 \approx t_2}$;
- if φ is of the form $t \in \mathrm{D}$ then prop_φ is $\mathsf{p}_{t \in \mathrm{D}}$;
- if φ is of the form $\neg\varphi_1$ then prop_φ is $\neg\mathsf{prop}_{\varphi_1}$;
- if φ is of the form $\varphi_1 \wedge \varphi_2$ then prop_φ is $\mathsf{prop}_{\varphi_1} \wedge \mathsf{prop}_{\varphi_2}$.

We also extend this propositional notation to probabilistic formulas: given a probabilistic formula δ of the form $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_\ell \cdot \mathsf{Pr}(\varphi_\ell) \bowtie q$ with $\bowtie \in \{\leq, \geq, <, >\}$, prop_δ represents the probabilistic propositional formula $q_1 \cdot \mathsf{Pr}(\mathsf{prop}_{\varphi_1}) + \cdots + q_\ell \cdot \mathsf{Pr}(\mathsf{prop}_{\varphi_\ell}) \bowtie q$.

Furthermore, we must import the algebraic requirements underlying the equational reasoning in the presence of domain restrictions to the propositional context. For this, assume that we want to test the satisfiability of $\delta \in \mathsf{Glob}$ and consider the set of relevant nominal terms for δ, $\mathsf{RelT}^\delta = \mathsf{subtrm}(\{\delta\} \cup \Delta^{\approx}_{\in}) \cup \{t{\downarrow} \mid t \in \mathsf{subtrm}(\{\delta\} \cup \Delta^{\approx}_{\in})\}$, where $\Delta^{\approx}_{\in} = \{\sigma(t) \approx \sigma(t') \mid (t \to t') \in R, \sigma \in \mathsf{subtrm}(\delta)^X\} \cup \{\sigma(t) \textcircled{\in} \mathrm{D} \mid (t \textcircled{\in} \mathrm{D}) \in \mathrm{RHS}, \sigma \in \mathsf{subtrm}(\delta)^X\}$ and $\mathrm{RHS} = \{t \textcircled{\in} \mathrm{D}'_1 \mid (t_1 \in \mathrm{D}_1, \ldots, t_{k_1} \in \mathrm{D}_{k_1} \Rightarrow t \textcircled{\in} \mathrm{D}'_1 \ldots, t \textcircled{\in} \mathrm{D}'_{k_2}) \in \Lambda\}$. RelT^δ incorporates the subterms of δ, their normal forms with respect to the convergent rewriting system R underlying Γ, and the equational theory and domain clauses instantiated on the subterms.

We achieve a sufficiently broad scope by defining the propositional symbols of interest as those that represent either equations between terms in RelT^δ or domain restrictions for such terms, which are gathered in the set $\mathcal{B}^\delta = \mathcal{B}^{\mathsf{Eq}} \cup \mathcal{B}^{\mathsf{DRes}}$, where $\mathcal{B}^{\mathsf{Eq}} = \left\{\mathsf{p}_{t_1 \approx t_2} \mid t_1, t_2 \in \mathsf{RelT}^\delta\right\}$ and $\mathcal{B}^{\mathsf{DRes}} = \left\{\mathsf{p}_{t \in \mathrm{D}} \mid t \in \mathsf{RelT}^\delta, \mathrm{D} \in \mathcal{D}\right\}$. Both equational statements and domain restrictions must obey some relations to be imposed on their representatives. These restrictions are established in Φ^δ, defined as follows:

$$\begin{aligned}
&\Phi^\delta = \{\mathsf{p}_{t \approx t{\downarrow}} \mid t \in \mathsf{RelT}^\delta\} \cup \{\mathsf{p}_{t_1 \approx t_2} \to \mathsf{p}_{t_2 \approx t_1} \mid t_1, t_2 \in \mathsf{RelT}^\delta\} \cup \{(\mathsf{p}_{t_1 \approx t_2} \wedge \mathsf{p}_{t_2 \approx t_3}) \to \mathsf{p}_{t_1 \approx t_3} \mid t_1, t_2, t_3 \in \mathsf{RelT}^\delta\} \cup \\
&\{(\mathsf{p}_{t_1 \approx t'_1} \wedge \ldots \wedge \mathsf{p}_{t_n \approx t'_n}) \to \mathsf{p}_{f(t_1,\ldots,t_n){\downarrow} \approx f(t'_1,\ldots,t'_n){\downarrow}} \mid t_1, t'_1, \ldots, t_n, t'_n, f(t_1,\ldots,t_n){\downarrow}, f(t'_1,\ldots,t'_n){\downarrow} \in \mathsf{RelT}^\delta\} \cup \\
&\{(\mathsf{p}_{t_1 \approx t_2} \wedge \mathsf{p}_{t_1 \in \mathrm{D}}) \to \mathsf{p}_{t_2 \in \mathrm{D}} \mid t_1, t_2 \in \mathsf{RelT}^\delta, \mathrm{D} \in \mathcal{D}\} \cup \\
&\{\textstyle\bigwedge_{i=1}^{k_1} \mathsf{p}_{\sigma(t_i){\downarrow} \in \mathrm{D}_i} \to \bigvee_{j=1}^{k_2} \mathsf{p}_{\sigma(t){\downarrow} \textcircled{\in} \mathrm{D}'_j} \mid \sigma \in (\mathsf{RelT}^\delta)^X, (t_1 \in \mathrm{D}_1, \ldots, t_{k_1} \in \mathrm{D}_{k_1} \Rightarrow t \textcircled{\in} \mathrm{D}'_1, \ldots, t \textcircled{\in} \mathrm{D}'_{k_2}) \in \Lambda\}.
\end{aligned} \tag{1}$$

We should emphasize that, since $\mathsf{subtrm}(\delta)$ has linear size on the length of δ and the equational theory is convergent, RelT^δ is well defined and has polynomial size on the length of δ. Denoting $\left|\mathsf{RelT}^\delta\right| = k$ and $|\mathcal{D}| = z$, Φ^δ has at most $k + k^2 + k^3 + k^{2a+2} + k^2 \cdot z + \lambda(k^{tmax} \cdot z^{dmax})$ elements, where a is the maximum arity of the function symbols occurring in RelT^δ, $|\Lambda| = \lambda$ and $tmax, dmax$ are the maximum number of terms and the maximum number of domain names occurring in a constraint in Λ. Sometimes we drop the superscript δ, provided that it is clear from context.

The subterm property provides control over the set Φ^δ, as long as it ensures that the domain restrictions over a term in RelT is only conditioned by domain restrictions over terms certainly in RelT. Thus, elements in Φ^δ are the necessary to reason about the domain restrictions that influence δ.

CNFSAT-DEqPrL problem: The CNFSAT-DEqPrL problem consists in deciding the existence of a model for a global formula $\delta \in \mathsf{Glob}$ given in conjunctive normal form. We analyze the CNFSAT-DEqPrL problem inspired on the developments for GenPSAT presented in (CCM17b) and explore a polynomial reduction to Satisfiability Modulo Theories with respect to the theory of quantifier-free linear arithmetic over the integers and reals ($\mathsf{QF\_LIRA}$) (BDEK07).

Assume that we are given a global formula $\delta \in \mathsf{Glob}$ given in CNF. Since each conjunct of δ is a disjunction of literals in $\forall\mathsf{Loc} \cup \neg\forall\mathsf{Loc} \cup \mathsf{Prob} \cup \neg\mathsf{Prob}$, we can rewrite it as: $\bigwedge_{j=1}^{m}(\forall\psi_1^j \vee \ldots \vee \forall\psi_{n_j}^j \vee \neg\forall\varphi_1^j \vee \ldots \vee \neg\forall\varphi_{k_j}^j \vee \xi_1^j \vee \ldots \vee \xi_{s_j}^j)$, where, for each $r \in \{1,\ldots,s_j\}$, the probabilistic literal ξ_r^j is assumed to take the following form: $q_{(r,j,1)} \cdot \mathsf{Pr}(\varphi_{(r,j,1)}) + \cdots + q_{(r,j,\ell_r^j)} \cdot \mathsf{Pr}(\varphi_{(r,j,\ell_r^j)}) \bowtie_r^j q_{(r,j)}$, with $\bowtie_r^j \in \{\geq,<\}$.

To address the need of witnesses for existential literals we need, at least, as many copies of $\mathcal{B}^\delta$ as the number of existential formulas $\neg\forall\mathsf{Loc}$ occurring in δ. In its description, δ counts with $\sum_{j=1}^m k_j$ literals of $\neg\forall\mathsf{Loc}$, so the final set of propositional symbols should contain all the required copies: $\bigcup_{j'=1}^m \bigcup_{\ell'=1}^{k_{j'}} \{\mathsf{p}^{[j',\ell']} \mid \mathsf{p} \in \mathcal{B}^\delta\}$. But, as we know, the probabilistic feature envisage a probability distribution over the set of valuations. In this sense, we should not limit our valuations to strictly represent witnesses for existential literals. Hence, we further need to consider an extra copy of $\mathcal{B}^\delta$.

Satisfying an element of the form $\forall\varphi$ imposes that φ must be verified in all possible outcome, whereas satisfying a formula as $\neg\forall\varphi$ requires that at least one possible outcome satisfies $\neg\varphi$. Therefore, our reduction to the propositional context must carry this sensitivity. In this way, the satisfiability of those literals is tested using several labeled copies of propositional variables (one copy for each literal of the form $\neg\forall\mathsf{Loc}$ plus the original copy), as if they had embedded several valuations. The *labels* are extended from the propositional variables to the propositional formulas as expected.

Prompted by the inclusion of SAT in PSAT (FDB11), $\mathsf{GenPSAT}$ (CCM17b), and $\mathsf{GGenPSAT}$ (CCM17a), the satisfiability of propositional formulas (representing literals in $\forall\mathsf{Loc}$) is tested by assigning them probability 1. Accordingly, and inspired on the $\mathsf{GenPSAT}$ normal forms (see (CCM17b)), we realize that the probabilistic (propositional) formulas to be tested should be *atomic*. For this purpose, we shall replace the propositional formulas occurring inside probabilistic (propositional) formulas by *ghost* propositional symbols. The existential literals are not supposed to influence probabilities (they have their own witnesses), so we discard them for a moment. Let us collect in $\mathfrak{GLoc}$ all the appropriate local formulas, suggested by δ, $\mathfrak{GLoc} = \bigcup_{j=1}^m \left(\{\psi_1^j,\ldots,\psi_{n_j}^j\} \cup \bigcup_{r=1}^{s_j}\{\varphi_{(r,j,1)},\ldots,\varphi_{(r,j,\ell_r^j)}\}\right)$, and in $\mathfrak{G}$ the corresponding propositional symbols: $\mathfrak{G} = \bigcup_{j=1}^m \left(\{\mathfrak{p}_{\psi_1^j},\ldots,\mathfrak{p}_{\psi_{n_j}^j}\} \cup \bigcup_{r=1}^{s_j}\{\mathfrak{p}_{\varphi_{(r,j,1)}},\ldots,\mathfrak{p}_{\varphi_{(r,j,\ell_r^j)}}\}\right)$. Furthermore, for each $\psi \in \mathfrak{GLoc}$, let the $[0,1]$-variable α_ψ represent the probability of ψ.

As we have already remarked, in order to obtain a correct translation into the propositional context, we should impose the requirements collected in Φ^δ (1). For this purpose, all the considered copies of $\mathcal{B}^\delta$ must verify those restrictions (with probability 1). And so, we should keep a special propositional *ghost* symbol for this purpose, $\mathfrak{p}_\phi$, and a variable to represent its probability, α_ϕ.

All these things considered, let $\widetilde{\mathcal{B}} = \mathcal{B}^\delta \cup \bigcup_{j'=1}^m \bigcup_{\ell'=1}^{k_{j'}} \{\mathsf{p}^{[j',\ell']} \mid \mathsf{p} \in \mathcal{B}^\delta\} \cup \mathfrak{G} \cup \{\mathfrak{p}_\phi\}$ represent the set of propositional symbols for our problem and denote by M the number of elements of $\mathfrak{G}\cup\{\mathfrak{p}_\phi\}$, $M \leq \sum_{j=1}^m \left(n_j + \sum_{r=1}^{s_j} \ell_r^j\right)+1$. For ease of notation, let $\nu : \mathfrak{GLoc}\cup\{\phi\} \to \{1,\ldots,M\}$

represent a bijection from the $\mathfrak{SLoc}$ coupled with the symbol ϕ to the set $\{1,\dots,M\}$ such that $\nu(\phi) = M$. The inverse bijection ν^{-1} is such that $\nu^{-1}(\{1,\dots,M\}) = \mathfrak{SLoc} \cup \{\phi\}$.

Let $H = [h_{ij}]$ denote a (still unknown) matrix of size $M \times (M+1)$ whose columns represent the valuations over $\widetilde{\mathcal{B}}$ evaluated on each propositional (ghost) symbol of $\mathfrak{G} \cup \{\mathfrak{p}_\phi\}$, i.e., $h_{ik} = v_k\left(\mathfrak{p}_{\nu^{-1}(i)}\right)$ for each $1 \le i \le M$ and $1 \le k \le M+1$. The $(M+1)$-vector $\pi = [\pi_k]$ represents a probability distribution over $\{v_1,\dots,v_{M+1}\}$. As we already mentioned, α_ψ represents the probability of each $\psi \in \mathfrak{SLoc}$ and α_ϕ represents the probability of Φ^δ.

To model all the possible valuations $\{v_1,...,v_{M+1}\}$, we consider $M+1$ copies of $\widetilde{\mathcal{B}}$: $\bigcup_{k=1}^{M+1}\{{}^{(k)}\mathsf{p} \mid \mathsf{p} \in \widetilde{\mathcal{B}}\}$. The idea is to test the satisfiability of δ through the assertion:

(prob) $\bigwedge_{j=1}^{m}\left(\bigvee_{s=1}^{n_j}\left(\alpha_{\psi_s^j} = 1\right) \vee \bigvee_{\ell=1}^{k_j}\left(\bigvee_{k=1}^{M+1} {}^{(k)}\mathsf{prop}^{[j,\ell]}_{\neg\varphi_\ell^j}\right) \vee \bigvee_{r=1}^{s_j}\left(\sum_{s=1}^{\ell_r^j} q_{(r,j,s)}\alpha_{\varphi_{(r,j,s)}} \bowtie_r^j q_{(r,j)}\right)\right)$,

subject to the additional assertions:

(prop_pos) $\bigwedge_{k=1}^{M+1}\left({}^{(k)}\mathfrak{p}_{\psi_s^j} \leftrightarrow \left(\bigwedge_{j'=1}^{m}\bigwedge_{\ell'=1}^{k_{j'}} {}^{(k)}\mathsf{prop}^{[j',\ell']}_{\psi_s^j} \wedge {}^{(k)}\mathsf{prop}_{\psi_s^j}\right)\right)$, for each $\psi_s^j \in \mathfrak{SLoc}$;

(prop_prob) $\bigwedge_{k=1}^{M+1}\left({}^{(k)}\mathfrak{p}_{\varphi_{(r,j,s)}} \leftrightarrow {}^{(k)}\mathsf{prop}_{\varphi_{(r,j,s)}}\right)$, for each $\varphi_{(r,j,s)} \in \mathfrak{SLoc}$;

(prop_phi) $\bigwedge_{k=1}^{M+1}\left({}^{(k)}\mathfrak{p}_\phi \leftrightarrow \bigwedge_{\phi\in\Phi^\delta}\left(\bigwedge_{j'=1}^{m}\bigwedge_{\ell'=1}^{k_{j'}} {}^{(k)}\phi^{[j',\ell']} \wedge {}^{(k)}\phi\right)\right)$;

(prob_phi) $\alpha_\phi = 1$;

(val1) $\left(\sum_{k=1}^{M+1} b_{ik} = \alpha_{\nu^{-1}(i)}\right)$, for each $i \in \{1,\dots,M\}$;

(val2) $((0 \le b_{ik} \le h_{ik}) \wedge (h_{ik} - 1 + \pi_k \le b_{ik} \le \pi_k))$, for $i \in \{1,...,M\}$ and $k \in \{1,...,M+1\}$;

(cons) $\left(h_{ik} = 1 \leftrightarrow {}^{(k)}\mathfrak{p}_{\nu^{-1}(i)}\right)$, for each $i \in \{1,\dots,M\}$ and $k \in \{1,\dots,M+1\}$;

(sums1) $\left(\sum_{i=1}^{M+1}\pi_k = 1\right)$.

So far we have introduced $\mathcal{O}(M + M \times (M+1))$ assertions, each of polynomial size on the length of δ, over a polynomial number of real, binary and propositional variables. It is easy to see that the presented translation to QF_LIRA is polynomial.

We test the satisfiability of δ by translating it into a QF_LIRA problem and then solving the latter appropriately. The procedure is presented in Algorithm 1. The procedure consists in initializing an empty QF_LIRA problem and then use the following auxiliary procedures: assert introduces an assertion into the QF_LIRA problem; lira_solver returns Sat or Unsat depending on whether the problem is satisfiable or not. When the resulting QF_LIRA problem is satisfiable, we conclude that δ is also satisfiable.

For the sake of illustration, we now use Algorithm 1 to decide whether a global formula is satisfiable or not.

Example 4.1. Recall Example 3.1 and consider the signature $\mathrm{F}^{\mathrm{xor}}$, the equational theory Γ^{xor} and the axiomatization Λ^{xor}. Let us test the satisfiability of the CNF global formula:

$$\mathsf{Pr}(n \approx \mathsf{zero}) \le \frac{2}{3}\cdot\mathsf{Pr}(n \in \mathsf{even}) \wedge \forall(n \in \mathsf{even}) \wedge \left(\neg\mathsf{Pr}(n \approx \mathsf{zero}) \le \frac{2}{3} \vee \neg\forall\mathsf{suc}(n) \in \mathsf{odd}\right),$$

with $n \in N$. We start by noting that $\mathsf{RelT}^\delta = \{n, \mathsf{zero}, \mathsf{suc}(n)\}$ and defining Φ^δ. Note also that $\mathfrak{SLoc} = \{n \approx \mathsf{zero}, n \in \mathsf{even}\}$, and consider the bijection $\nu : \mathfrak{SLoc} \cup \{\phi\} \to \{1,2,3\}$ such that $\nu(n \approx \mathsf{zero}) = 1$, $\nu(n \in \mathsf{even}) = 2$, $\nu(\phi) = 3$.

Algorithm 1 $\mathsf{CNFSAT\text{-}DEqPrL}$ solver based on $\mathsf{SMT - QF\_LIRA}$

procedure CNFSATDEqPrL

  input: CNF formula δ: $\bigwedge_{j=1}^{m} \left(\forall\psi_1^j \vee \ldots \vee \forall\psi_{n_j}^j \vee \neg\forall\varphi_1^j \vee \ldots \vee \neg\forall\varphi_{k_j}^j \vee \xi_1^j \vee \ldots \vee \xi_{s_j}^j\right)$

  output: Sat or Unsat depending on whether δ is satisfiable or not

  assume: $M := \sum_{j=1}^{m} (n_j + \sum_{r=1}^{s_j} \ell_r^j) + 1$

    $\nu : \mathfrak{SLoc} \cup \{\phi\} \to \{1, \ldots, M\}$ is a bijection

  declare: prop. variables: $\bigcup_{k=1}^{M+1}\left({}^{(k)}\mathcal{B}^\delta \cup \bigcup_{j'=1}^{m} \bigcup_{\ell'=1}^{k_{j'}} \{{}^{(k)}\mathsf{p}^{[j',\ell']} \mid \mathsf{p} \in \mathcal{B}^\delta\} \cup {}^{(k)}\mathfrak{G} \cup \{{}^{(k)}\mathfrak{p}_\phi\}\right)$

    binary variables: h_{ik}, for $i \in \{1, \ldots, M\}$, $k \in \{1, \ldots, M+1\}$

    $[0,1]$-variables: $\alpha_{\nu^{-1}(i)}, \pi_k, b_{ik}$, for $i \in \{1, \ldots, M\}$, $k \in \{1, \ldots, M+1\}$

  for $j = 1$ to m **do**

    $\mathsf{assert}\left(\bigwedge_{k=1}^{M+1} \bigwedge_{s=1}^{n_j} \left({}^{(k)}\mathfrak{p}_{\psi_s^j} \leftrightarrow \left(\bigwedge_{j'=1}^{m} \bigwedge_{\ell'=1}^{k_{j'}} {}^{(k)}\mathsf{prop}_{\psi_s^j}^{[j',\ell']} \wedge {}^{(k)}\mathsf{prop}_{\psi_s^j}\right)\right)\right)$ ▷ (prop_pos)

    $\mathsf{assert}\left(\bigwedge_{k=1}^{M+1} \bigwedge_{r=1}^{s_j} \bigwedge_{s=1}^{\ell_r^j} \left({}^{(k)}\mathfrak{p}_{\varphi_{(r,j,s)}} \leftrightarrow {}^{(k)}\mathsf{prop}_{\varphi_{(r,j,s)}}\right)\right)$ ▷ (prop_prob)

  for $i = 1$ to M **do**

    $\mathsf{assert}\left(\sum_{k=1}^{M+1} b_{ik} = \alpha_{\nu^{-1}(i)}\right)$ ▷ (val1)

    for $k = 1$ to $M+1$ **do**

      $\mathsf{assert}\left((0 \leq b_{ik} \leq h_{ik}) \wedge (h_{ik} - 1 + \pi_k \leq b_{ik} \leq \pi_k)\right)$ ▷ (val2)

      $\mathsf{assert}\left(h_{ik} = 1 \leftrightarrow {}^{(k)}\mathfrak{p}_{\nu^{-1}(i)}\right)$ ▷ (cons)

  $\mathsf{assert}\left(\bigwedge_{j=1}^{m}\left(\bigvee_{s=1}^{n_j}\left(\alpha_{\psi_s^j}=1\right) \vee \bigvee_{\ell=1}^{k_j}\left(\bigvee_{k=1}^{M+1} {}^{(k)}\mathsf{prop}_{\neg\varphi_\ell^j}^{[j,\ell]}\right) \vee \bigvee_{r=1}^{s_j}\left(\sum_{s=1}^{\ell_r^j} q_{(r,j,s)}\alpha_{\varphi_{(r,j,s)}} \bowtie_r^j q_{(r,j)}\right)\right)\right)$ ▷(prob)

  $\mathsf{assert}\left(\bigwedge_{k=1}^{M+1}\left({}^{(k)}\mathfrak{p}_\phi \leftrightarrow \bigwedge_{\phi \in \Phi^\delta}\left(\bigwedge_{j'=1}^{m} \bigwedge_{\ell'=1}^{k_{j'}} {}^{(k)}\phi^{[j',\ell']} \wedge {}^{(k)}\phi\right)\right)\right)$ ▷ (prop_phi)

  $\mathsf{assert}\,(\alpha_\phi = 1)$ ▷ (prob_phi)

  $\mathsf{assert}\left(\sum_{k=1}^{M+1} \pi_k = 1\right)$ ▷ (sums1)

  return lira_solver() ▷return Sat if the assertions are satisfiable, Unsat otherwise

For the given formula, the assertion (prob) reads like

$$\left(\alpha_{n\approx\mathsf{zero}} \leq \frac{2}{3} \cdot \alpha_{n\in\mathsf{even}}\right) \wedge (\alpha_{n\in\mathsf{even}} = 1) \wedge \left(\alpha_{n\approx zero} > \frac{2}{3} \;\vee\; \bigvee_{k=1}^{4} {}^{(k)}\mathsf{prop}_{\neg\mathsf{suc}(n)\in\mathsf{odd}}^{[3,1]}\right),$$

which together with the remaining assertions carefully described in Algorithm 1 is unsatisfiable. To check that, assume that it would have a solution (denoted by x^* for each variable x) and let us derive a contradiction.

Begin noting that by (val1), $b_{\nu^{-1}(n\in\mathsf{even}),k}$ ranges in the interval $[0,\pi_k]$ for each $k\in\{1,2,3\}$. Once $\alpha^*_{n\in\mathsf{even}}=1$, then every $b_{\nu^{-1}(n\in\mathsf{even}),k}$ should coincide with π_k and, by (val2), $h^*_{\nu^{-1}(n\in\mathsf{even}),k}=1$ for every $k\in\{1,2,3\}$. Then, by (cons), ${}^{(k)}\mathfrak{p}_{n\in\mathsf{even}}$ holds. But, by (prop_pos) this means that for each k, ${}^{(k)}\mathsf{prop}^{[3,1]}_{n\in\mathsf{even}}\wedge{}^{(k)}\mathsf{prop}_{n\in\mathsf{even}}$ also holds. Observing that $(n\in\mathsf{even}\to\mathsf{suc}(n)\in\mathsf{odd})\in\Phi^\delta$, it follows that for each k, ${}^{(k)}\mathsf{prop}^{[3,1]}_{\mathsf{suc}(n)\in\mathsf{odd}}\wedge{}^{(k)}\mathsf{prop}_{\mathsf{suc}(n)\in\mathsf{odd}}$ holds. Then, we have that ${}^{(k)}\mathsf{prop}^{[3,1]}_{\neg\mathsf{suc}(n)\in\mathsf{odd}}$ does not hold for every k. On the other hand, since $\alpha^*_{n\approx\mathsf{zero}}\le\frac{2}{3}$, there is no way for the last conjunct to hold and we conclude that the formula is unsatisfiable. △

Now that we checked how to apply the procedure, let us state its correctness (see a sketch of the proof in the Appendix and the details in (Mor17)).

Lemma 4.1. If Γ is a convergent equational theory and Λ is a set of domain clauses with the subterm property, a global formula $\delta\in\mathsf{Glob}$ in CNF is satisfiable iff Algorithm 1 returns Sat.

Tseitin-like transformation on DEqPrL: So far, we have described an algorithm to decide the satisfiability of a global formulas in CNF. However, transforming a global formula into CNF eventually leads to an explosion in the length of the formula. Luckily, we have a Tseitin-like transformation for DEqPrL, which provides a method to transform any global formula into an equisatisfiable CNF formula with linear size on the length of the original formula, and allows us to take advantage of the CNFSAT-DEqPrL solver.

The idea is to introduce additional atoms $\forall(n_1^{\delta'}\approx n_2^{\delta'})$ for every non-atomic subformula δ' of δ, ensure that $\forall(n_1^{\delta'}\approx n_2^{\delta'})\leftrightarrow\delta'$ and, in the end, additionally ensure that the former formula is satisfied by imposing $\forall(n_1^{\delta}\approx n_2^{\delta})$. In this sense, given a global formula $\delta\in\mathsf{Glob}$, we consider the set of all subformulas of δ that are not atoms, $\mathsf{subform}(\delta)\setminus(\forall\mathsf{Loc}\cup\mathsf{Prob})$, and fix a pair of new (and distinct) names for each of them. To ease notation, we denote by $\mathsf{GA}(\delta')$ the atom corresponding to the subformula $\delta'\in(\mathsf{subform}(\delta)\setminus(\forall\mathsf{Loc}\cup\mathsf{Prob}))$. Furthermore, we abuse notation and also denote an atom $\delta'\in(\mathsf{subform}(\delta)\cap(\forall\mathsf{Loc}\cup\mathsf{Prob}))$ by $\mathsf{GA}(\delta')$. In short,

$$\mathsf{GA}(\delta')=\begin{cases}\delta' & \text{if } \delta'\in(\forall\mathsf{Loc}\cup\mathsf{Prob})\\ \forall(n_1^{\delta'}\approx n_2^{\delta'}) & \text{otherwise}\end{cases}$$

For each subformula $\delta'\in(\mathsf{subform}(\delta)\setminus(\forall\mathsf{Loc}\cup\mathsf{Prob}))$, we define the additional conjuncts $\mathsf{tc}(\delta')$ representing the equivalence $\mathsf{GA}(\delta')\leftrightarrow\delta'$ in CNF according to the structure of δ':

$\mathsf{tc}(\neg\psi)=(\mathsf{GA}(\neg\psi)\vee\mathsf{GA}(\psi))\wedge(\neg\mathsf{GA}(\neg\psi)\vee\neg\mathsf{GA}(\psi))$;

$\mathsf{tc}(\psi_1\wedge\psi_2)=(\neg\mathsf{GA}(\psi_1\wedge\psi_2)\vee\mathsf{GA}(\psi_1))\wedge(\neg\mathsf{GA}(\psi_1\wedge\psi_2)\vee\mathsf{GA}(\psi_2))\wedge(\mathsf{GA}(\psi_1\wedge\psi_2)\vee\neg\mathsf{GA}(\psi_1)\vee\neg\mathsf{GA}(\psi_2))$;

$\mathsf{tc}(\psi_1\vee\psi_2)=(\mathsf{GA}(\psi_1\vee\psi_2)\vee\neg\mathsf{GA}(\psi_1))\wedge(\mathsf{GA}(\psi_1\vee\psi_2)\vee\neg\mathsf{GA}(\psi_2))\wedge(\neg\mathsf{GA}(\psi_1\vee\psi_2)\vee\mathsf{GA}(\psi_1)\vee\mathsf{GA}(\psi_2))$.

We define the Tseitin-like transformation on DEqPrL simply as:

$$\mathsf{tt}(\delta)=\mathsf{GA}(\delta)\wedge\bigwedge_{\delta'\in(\mathsf{subform}(\delta)\setminus(\forall\mathsf{Loc}\cup\mathsf{Prob}))}\mathsf{tc}(\delta').$$

Notice that the obtained CNF formula has linear size on the length of δ, since

$\mathsf{subform}(\delta)$ has linear size on the length of δ and the transformation $\mathsf{tc}(\cdot)$ increments the length of the formula only by a constant. As a corollary of the previous construction we have the following Lemma.

Lemma 4.2. Given $\delta \in \mathsf{Glob}$, there exists an equisatisfiable formula $\delta' \in \mathsf{Glob}$ in conjunctive normal form whose length is linear on the length of δ.

SAT-DEqPrL **problem:** In general, we are looking for a procedure to decide SAT-DEqPrL. Fortunately, the Tseitin-like transformation for DEqPrL and the CNFSAT-DEqPrL solver will greatly ease our task. Given a global formula $\delta \in \mathsf{Glob}$, we seek out an equisatisfiable formula δ' in CNF and then use the CNFSAT-DEqPrL solver to decide about the existence of a model for δ' (and for δ).

Theorem 4.1. If Γ is a convergent equational theory and Λ is a set of domain clauses with the subterm property, then the SAT-DEqPrL problem is decidable.

Proof. Given $\delta \in \mathsf{Glob}$, we use the Tseitin-like transformation for DEqPrL to convert δ into an equisatisfiable formula $\mathsf{tt}(\delta)$ in conjunctive normal form. Then, we run the CNFSAT-DEqPrL solver presented in Algorithm 1 on $\mathsf{tt}(\delta)$. If CNFSAT-DEqPrL returns Sat then $\mathsf{tt}(\delta)$ has a model, and so δ has a model; otherwise δ will be unsatisfiable. □

4.2. *Validity*

The decidability of the logic follows as an immediate corollary of the satisfiability result.

Theorem 4.2. If Γ is a convergent equational theory and Λ is a set of domain clauses with the subterm property, then the logic is decidable.

Proof. Since the deduction metatheorem holds, given a finite set $\Delta \subseteq \mathsf{Glob}$ and $\varphi \in \mathsf{Glob}$, proving $\Delta \vdash_{(\Gamma,\Lambda)} \varphi$ is equivalent to proving $\vdash_{(\Gamma,\Lambda)} ((\bigwedge_{\psi\in\Delta} \psi) \to \varphi)$, so we proceed by checking the validity problem. Given $\delta \in \mathsf{Glob}$, we decide whether $\vdash_{(\Gamma,\Lambda)}\delta$ or $\nvdash_{(\Gamma,\Lambda)}\delta$ by testing the satisfiability of $\neg\delta$: if $\neg\delta$ is satisfiable, since the logic is sound, we conclude that $\nvdash_{(\Gamma,\Lambda)} \delta$; if $\neg\delta$ is not satisfiable, we use completeness to conclude that $\vdash_{(\Gamma,\Lambda)} \delta$. □

4.3. *Complexity*

The satisfiability result highlights a way of deciding SAT-DEqPrL by reduction to a QF_LIRA solver, under the assumption that Γ is a convergent equational theory and Λ is a set of domain clauses with the subterm property. We will now analyse complexity of the procedures previously obtained.

As we already observed, the CNFSAT-DEqPrL solver presented in Algorithm 1 exhibits a polynomial reduction from CNFSAT-DEqPrL to QF_LIRA. The complexity result for the satisfiability problem CNFSAT-DEqPrL is parametric and also depends on the complexity of determining normal forms for terms with respect to the equational specification of the algebraic basis, which are fundamental to obtain the set of relevant terms RelT^{δ}. The complexity of CNFSAT-DEqPrL is the same as for QF_LIRA as long as the complexity of computing normal forms with respect to Γ (dub it the $\Gamma\!\downarrow$-*problem*) is in P.

Corollary 4.1. Assuming that Γ is a convergent equational theory whose $\Gamma\!\downarrow$-problem is in P and Λ is a set of domain clauses with the subterm property, then the satisfiability problem $\mathsf{CNFSAT\text{-}DEqPrL}$ is in NP and the validity problem for DNF formulas in DEqPrL is in coNP.

Note that when the rewriting system underlying the equational theory Γ is subterm convergent, the complexity class of the $\Gamma\!\downarrow$-problem is in P. We should also remark that SAT can obviously be modeled in DEqPrL, by assigning an atom $\forall(n_1 \approx n_2)$ composed by two fresh names n_1, n_2 to each propositional symbol to be considered.

Corollary 4.2. If Γ is a subterm theory and Λ is a set of domain clauses with the subterm property, then $\mathsf{CNFSAT\text{-}DEqPrL}$ is NP-complete.

The complexity of the satisfiability problem for the logic is now immediate from the complexity of $\mathsf{CNFSAT\text{-}DEqPrL}$ and by Lemma 4.2.

Corollary 4.3. Assuming that Γ is a convergent equational theory whose $\Gamma\!\downarrow$-problem is in P and Λ is a set of domain clauses with the subterm property, then the satisfiability problem $\mathsf{SAT\text{-}DEqPrL}$ is in NP and the validity problem for DEqPrL is in coNP.

Corollary 4.4. If the equational theory of Γ is generated by a subterm convergent rewriting system and Λ is a set of domain clauses with the subterm property, then the $\mathsf{SAT\text{-}DEqPrL}$ problem is NP-complete.

5. Examples

Now we model some information security examples in DEqPrL and observe how important are the implementation details on the estimation of probabilities of the existence of attacks to cryptographic protocols.

5.1. *Offline Guessing Attacks with some Cryptanalysis*

Now we focus on the analysis of offline guessing attacks to cryptographic protocols-(Bau05) in the context of DEqPrL. Actually, we may focus in a wider and more expressive formulation where the attacker, besides all the algebraic knowledge he has about the protocol and cryptographic primitives, is also endowed with some cryptanalytic capabilities. To analyze offline guessing one assumes that the attacker observed messages named $m_1, \ldots, m_k$ which were built as $t_1, \ldots, t_k \in T(N)$, but the attacker cannot know the concrete values of the random and secret names used to build them. Still, he can try to mount an attack by guessing some secrets $s_1, \ldots, s_n \in N$ used by the parties executing the protocol. The attack is successful if the attacker can distinguish whether his guesses $s_1^*, \ldots, s_n^* \in N$ are correct or not.

Definition 5.1. Let $m_1, \ldots, m_k \in T(N)$ represent the messages exchanged by the parties executing a given cryptographic protocol, and Γ denote the equational specification of the underlying algebraic basis and Λ collects the domain restrictions on terms. The protocol is susceptible to an *offline guessing attack using cryptanalysis* if there exists a *recipe* $\varphi \in \mathsf{Loc}$, with $\mathsf{subtrm}(\varphi) \subseteq T(\{m_1, \ldots, m_k, s_1^*, \ldots, s_n^*\})$ such that:

$$\begin{array}{lccc} & \forall(m_1 \approx t_1 \wedge \cdots \wedge m_k \approx t_k) & \nvdash_{(\Gamma,\Lambda)} & \forall\varphi \\ \text{and} & & & \\ & \forall(m_1 \approx t_1 \wedge \cdots \wedge m_k \approx t_k) & \vdash_{(\Gamma,\Lambda)} & \forall(s_1^* \approx s_1 \wedge \cdots \wedge s_n^* \approx s_n \to \varphi)\,. \end{array}$$

Note that the recipe is a formula involving equations and domain restrictions and is constructed exclusively from messages observed by the attacker and from guesses for the secret values. The recipe should not be derivable in general, but should be valid under the assumption that the attacker correctly guessed the secrets, proving to constitute a reliable formula for the attacker to check whether he actually guessed the secrets.

This task is undecidable in general as the recipe may be arbitrarily complex, but for subterm convergent rewriting systems the problem is decidable, as only a finite number of 'dangerous' recipes need to be tested (AC05; AC06; Bau05).

The analysis of the existence of offline guessing attacks is even more interesting when probabilities come into play, as the attacker will be able to narrow the set of possible secrets. In these lines, under appropriate probabilistic conditions and applying axiom P3, one should be able to estimate the probability of offline guessing attacks in DEqPrL.

Example 5.1. As an application, consider a protocol adapted from (CE04), where $a, b, n_a, p_{ab} \in N$:

$$1.\ a \to b : (a, n_a)$$

$$2.\ b \to a : \{n_a\}_{p_{ab}}$$

In the first step, some party named a sends a message to another party named b in order to initiate some communication session. The message is a pair containing a's name and a random value (*nonce*) named n_a, that a generated freshly, and which is intended to distinguish this request from other, similar, past or future, requests. Upon reception of the first message, b responds by ciphering n_a with a secret password p_{ab} shared with a. When receiving the second message, a can decrypt it and recognize b's response to his request to initiate a session.

In this case, it is simple to observe that the secret shared password p_{ab} is vulnerable to an offline guessing attack. Suppose that the attacker observes the execution of the protocol by parties a and b, and got hold of the two exchanged messages m_1 and m_2. He can now manipulate these messages, using his guess p^*_{ab} of p_{ab}, and come up with the recipe $\{m_2\}^{-1}_{p^*_{ab}} \approx \pi_2(m_1)$. Indeed, only under the correct guess we can prove that the decryption of m_2 with p^*_{ab} coincides with the second projection of m_1, that is, n_a. We can use our logic and, in particular, axioms $\mathrm{E}(\Gamma^{\mathsf{DY}})$ that encode the equations in Γ^{DY} to check that, indeed,

$$\forall (m_1 \approx (a, n_a) \wedge m_2 \approx \{n_a\}_{p_{ab}}) \nvdash_{(\Gamma^{\mathsf{DY}}, \Lambda^{\mathsf{DY}})} \forall (\{m_2\}^{-1}_{p^*_{ab}} \approx \pi_2(m_1)) \text{ and}$$

$$\forall (m_1 \approx (a, n_a) \wedge m_2 \approx \{n_a\}_{p_{ab}}) \vdash_{(\Gamma^{\mathsf{DY}}, \Lambda^{\mathsf{DY}})} \forall (p^*_{ab} \approx p_{ab} \to \{m_2\}^{-1}_{p^*_{ab}} \approx \pi_2(m_1)),$$

The existence of an offline guessing attack for this protocol led to an improvement of the exchanged messages by concatenating a confounder c with the nonce and encrypting with the public key $\mathsf{pub}(b)$ afterwards, giving rise to Gong's protocol (GLNS93):

$$1.\ a \to b : \{\!|(n_a, c)|\!\}_{\mathsf{pub}(b)}$$

$$2.\ b \to a : \{n_a\}_{p_{ab}}\ .$$

Gong's protocol was proved to be secure against offline guessing (CE04; GLNS93), in the sense that the probability of an attack is negligible. We will observe that such security highly depends on the practical implementation of the protocol. This is one of the great achievements that we obtain with DEQPRL: we are able to cover some implementation details formally within the logic and conclude how do they compromise security.

Let us extend the set of domain names $\mathcal{D} = \mathcal{D}^{\mathsf{DY}} \cup \{\mathsf{conf}\}$ and, further, assume that the confounder c is sampled uniformly from a set with M elements, and that the set of symmetric keys from which p_{ab} is uniformly chosen has N elements. The estimation of the probability of an offline guessing attack on the independent names p_{ab} and c, with guesses p^*_{ab} and c^*, is given by:

$$\mathsf{Hyp} \vdash_{(\Gamma^{\mathsf{DY}},\Lambda^{\mathsf{DY}})} \mathsf{Pr}(p_{ab} \approx p^*_{ab} \wedge c \approx c^*) \leq \mathsf{Pr}(\{\!|(\{m_2\}^{-1}_{p^*_{ab}}, c^*)|\!\}_{\mathsf{pub}(b)} \approx m_1) \ ,$$

where the set of hypothesis consists of the uniform probabilities and independence of p^*_{ab} and c^*, of a record of the exchanged messages and of some cryptanalytic properties,

$$\mathsf{Hyp} = \{\forall(c^* \in \mathsf{conf}) \to \mathsf{Pr}(c \approx c^*) = \tfrac{1}{M},\ \forall(p^*_{ab} \in \mathsf{sym\_key}) \to \mathsf{Pr}(p_{ab} \approx p^*_{ab}) = \tfrac{1}{N},\ \mathsf{Ind}^{p^*_{ab},c^*}_{N,M},$$
$$\forall(c^* \in \mathsf{conf}),\ \forall(p^*_{ab} \in \mathsf{sym\_key}),\ \forall(m_1 \approx \{\!|(n_a, c)|\!\}_{\mathsf{pub}(b)}\ \wedge\ m_2 \approx \{n_a\}_{p_{ab}})\}\,.$$

According to the independence property for p^*_{ab} and c^*, the probability of guessing c and p_{ab}, and therefore the probability of success of an offline guessing attack is given by

$$\mathsf{Hyp} \vdash_{(\Gamma^{\mathsf{DY}},\Lambda^{\mathsf{DY}})} \frac{1}{N \cdot M} \ \leq\ \mathsf{Pr}(\{\!|(\{m_2\}^{-1}_{p^*_{ab}}, c^*)|\!\}_{\mathsf{pub}(b)} \approx m_1)\ .$$

Often, symmetric keys are defined as being *weak keys*, meaning that they are chosen from small sample spaces. In this sense, N is usually small. On the contrary, the commonly called *unguessable* values are believed to be chosen from very big sets. However, in the practical implementation of protocols it does not always happen, and we can model it in our logic. Notice that if M is also a small number, the probability of an attack is not negligible, as it is minimized by the non-negligible value $\frac{1}{N \cdot M}$. △

5.2. *On the Implementation Details*

The reduced range of values taken by some critical parameters in the concrete implementation of cryptographic protocols can seriously compromise their security. Recently (see (ABD+15)) it was shown that some modern implementations of Diffie-Hellman key exchange are vulnerable to attacks from adversaries with reasonable resources.

A Diffie-Hellman key exchange consists of a preliminary agreement of a large prime p and a generator g by agents a and b, then both parties generate random integers x_a and x_b. Once all the values are fixed, a sends the exponentiation of g with x_a modulo p to b, and b sends the exponentiation of g with its private key x_b modulo p to a. At the end of the protocol, a and b are sharing the secret $(g^{x_a})^{x_b} \mathsf{mod}\ p = (g^{x_b})^{x_a} \mathsf{mod}\ p$. Computing discrete logarithms remains the best known cryptanalytic attack to the security of Diffie-Hellman. In general, discrete log computations for arbitrary primes are known to take enough time to ensure that any session expires before the intruder carries out an attack, but Logjam (ABD+15) presents a technique that uses number field sieve and allows one to compute the discrete log of primes in a specified 512-bit group in about one minute, by means of a precomputation of the first three steps of number field sieve for that specific group. In fact, this vulnerability was already known since 1992 (BBDL+15), but was applied by Logjam (ABD+15) to downgrade a TLS connection to use 512-bit Diffie-Hellman export-grade cryptography, through a man-in-the-middle network attacker. Let us analyze formally, within DEqPrL, how would a cryptanalytic attack through the discrete log compromise the security of Diffie-Hellman.

Example 5.2. Consider a Diffie-Hellman key exchange protocol:

$$1.\ a \to b : g^{x_a} \ \mathsf{mod}\ p$$
$$2.\ b \to a : g^{x_b} \ \mathsf{mod}\ p$$

Let us assume the attacker possesses enough computational resources to manage a precomputation of the first steps of number field sieve for a chosen group of 512-bit prime. Recall that the discrete logs in that group are then computed in a feasible amount of time. So, we can consider, in our signature, a function symbol representing the discrete log for each of those primes. Consider the signature F^{DH} containing: $\mathsf{DLOG}_{(\cdot)}(\cdot,\cdot) \in \mathrm{F}_3^{\mathsf{DH}}$ representing an oracle for the discrete log of the subscript argument; $(\cdot)^{(\cdot)} \in \mathrm{F}_2^{\mathsf{DH}}$ representing exponentiation; $(\cdot)\ \mathsf{mod}\ (\cdot) \in \mathrm{F}_2^{\mathsf{DH}}$ representing the remainder of the division of the first by the second argument. In the context of Diffie-Hellman key exchange, the equational properties of these operations are given by: $\Gamma^{\mathsf{DH}} = \{((x^{x_1})^{x_2}\ \mathsf{mod}\ x_3) \approx ((x^{x_2})^{x_1}\ \mathsf{mod}\ x_3)\}$. Now let us fix some domains, representing the chosen group of 512-bit primes for the implementation, the set of generators, the set of private keys and the set of ciphertexts: $\mathcal{D}^{\mathsf{DH}} = \{\ 512\_\mathsf{prime},\ \mathsf{gen},\ \mathsf{prv\_key},\ \mathsf{ciphertxt}\ \}$. We axiomatize the domain restrictions simply as: $\Lambda^{\mathsf{DH}} = \{\ (x \in \mathsf{prv\_key}, g \in \mathsf{gen}, p \in 512\_\mathsf{prime} \Rightarrow (g^x\ \mathsf{mod}\ p) \in \mathsf{ciphertxt})\ \}$.

The probability of a cryptanalytic attack using discrete log can be expressed in DEqPrL as: $\mathsf{Hyp}^{\mathsf{DH}} \vdash_{(\Gamma^{\mathsf{DH}},\Lambda^{\mathsf{DH}})} \mathsf{Pr}(\mathsf{DLOG}_p(g,m_1) \approx x_a) \geq \mathsf{Pr}(p \in 512\_\mathsf{prime})$, where $\mathsf{Hyp}^{\mathsf{DH}} = \{\forall(m_1 \approx g^{x_a} \mathsf{mod}\ p \wedge m_2 \approx g^{x_b} \mathsf{mod}\ p), \forall(p \in 512\_\mathsf{prime} \to \mathsf{DLOG}_p(x_1, x_1^{x_2}\ \mathsf{mod}\ p) \approx x_2)\}$, meaning that the probability of an offline guessing attack is bounded below by the probability of the intruder's smart choice for the group to which he develops the precomputation actually fall within the choice of the implementer. Obviously, the attacker would not waste resources precomputing discrete logarithms unlikely to be used. There are groups of 512-bit primes known to be much popular than others, so the probability of the intruder's smart choice be within one of the implementer's choice can be significantly large, thereby influencing the probability of the existence of an attack. This formalization should be seen as a simple illustration of how the cryptanalytic attacks can be modeled within DEqPrL. △

6. Conclusion and future work

We combined aspects from classical, equational and probabilistic reasoning to construct a logic suited for the qualitative and quantitative analysis of equational constraints and domain restrictions over a set of outcomes. The design of the logic was aimed at formalizing the kind of reasoning carried out in security protocol analysis provided an attacker with cryptanalytic capabilities. Parameterized by suitable properties of the underlying algebraic base and domain restrictions, we have obtained a sound and weakly complete deductive system for our logic. We found a polynomial reduction from CNFSAT-DEqPrL to QF_LIRA, provided that the algebraic basis is given by means of a convergent rewriting system and, additionally, that the axiomatization of domain restrictions enjoys a suitable subterm property. The complexity result followed naturally from the previous constructions. The decidability result also took advantage of the way in which the strategy was conducted. Lastly, we used the logic to verify and estimate the probability of attacks to cryptographic protocols in the presence of an attacker with an informed way of cryptanalysis, reducing the gap between symbolic and computational models (Bau05; AC06; AC05; CBC10; CBC13). These results enabled the implementation of a prototype tool for the satisfiability problem of DEqPrL that can be found in (CCM16) but is out of the scope of this paper - for more details see (Mor17).

Even though our decidability results cover a very interesting range of examples, it would be interesting to explore their extension in order to handle decidable equational theories in general, i.e. not necessarily defined by means of convergent rewriting systems (DOS88).

References

D. Adrian, K. Bhargavan, Durumeric, et al. Imperfect forward secrecy: How Diffie-Hellman fails in practice. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pages 5–17. ACM, 2015.

M. Abadi and V. Cortier. Deciding knowledge in security protocols under (many more) equational theories. In *Proceedings of the 18th IEEE Computer Security Foundations Workshop (CSFW'05)*, pages 62–76. IEEE, 2005.

M. Abadi and V. Cortier. Deciding knowledge in security protocols under equational theories. *Theoretical Computer Science*, 367(1):2–32, 2006.

M. Baudet. Deciding security of protocols against off-line guessing attacks. In *Proceedings of 12th ACM Conference on Computer and Communications Security*, pages 16–25. ACM, 2005.

B. Beurdouche, K. Bhargavan, A. Delignat-Lavaud, C. Fournet, M. Kohlweiss, A. Pironti, P. Y. Strub, and J. K. Zinzindohoue. A messy state of the union: Taming the composite state machines of tls. In *Symposium on Security and Privacy 2015*, pages 535–552. IEEE, 2015.

B. Becker, C. Dax, J. Eisinger, and F. Klaedtke. LIRA: Handling constraints of linear arithmetics over the integers and the reals. In *International Conference on Computer Aided Verification*, volume 4590 of *Lecture Notes in Computer Science*, pages 307–310. Springer, 2007.

F. Baader and T. Nipkow. *Term rewriting and all that.* Cambridge University Press, 1999.

B. Conchinha, D. Basin, and C. Caleiro. Efficient decision procedures for message deducibility and static equivalence. In *Proceedings of 7th International Workshop on Formal Aspects in Security and Trust*, volume 6561 of *LNCS*, pages 34–49. Springer, 2010.

B. Conchinha, D. Basin, and C. Caleiro. Symbolic probabilistic analysis of off-line guessing. In *European Symposium on Research in Computer Security*, volume 8134 of *Lecture Notes in Computer Science*, pages 363–380. Springer, 2013.

C. Caleiro, F. Casal, and A. Mordido. DNFSAT-DEqPrL solver, 2016. Available online at https://github.com/fcasal/satdeqprl.git.

C. Caleiro, F. Casal, and A. Mordido. Classical generalized probabilistic satisfiability. In *International Joint Conf. on Artificial Intelligence - IJCAI*, August 2017.

Carlos Caleiro, Filipe Casal, and Andreia Mordido. Generalized probabilistic satisfiability. *Electronic Notes in Theoretical Computer Science*, 332:39–56, 2017.

R. J. Corin and S. Etalle. A simple procedure for finding guessing attacks. 2004.

V. Cortier, S. Kremer, and B. Warinschi. A survey of symbolic methods in computational analysis of cryptographic systems. *Journal of Automated Reasoning*, 46(3-4):225–259, 2011.

N. Dershowitz, M. Okada, and G. Sivakumar. Canonical conditional rewrite systems. In *Proceedings of the 9th International Conference on Automated Deduction*, volume 310 of *Lecture Notes in Computer Science*, pages 538–549. Springer, 1988.

D. Dolev and A. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–208, 1983.

M. Finger and G. De Bona. Probabilistic satisfiability: Logic-based algorithms and phase transition. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'11)*, pages 528–533, 2011.

R. Fagin, J. Y. Halpern, and N. Megiddo. A logic for reasoning about probabilities. *Information and Computation*, 87(1):78–128, 1990.

L. Gong, M. A. Lomas, R. M. Needham, and J. H. Saltzer. Protecting poorly chosen secrets from guessing attacks. *Journal on Selected Areas in Communications*, 11(5):648–656, 1993.

L. Henkin. The completeness of the first-order functional calculus. *The Journal of Symbolic Logic*, 14(03):159–166, 1949.

B. Montalto and C. Caleiro. Modeling and reasoning about an attacker with cryptanalytical capabilities. *Electronic Notes in Theoretical Computer Science*, 253(3):143–165, 2009.

A. Mordido and C. Caleiro. An equation-based classical logic. In *Proceedings of the 22nd International Workshop on Logic, Language, Information, and Computation*, volume 9160 of *Lecture Notes in Computer Science*, pages 38–52. Springer, 2015.

A. Mordido. *A probabilistic logic over equations and domain restrictions.* PhD Thesis, IST, Universidade de Lisboa, 2017.

P. Mateus, A. Sernadas, and C. Sernadas. Exogenous semantics approach to enriching logics. *Essays on the Foundations of Mathematics and Logic*, 1:165–194, 2005.

R. Nieuwenhuis, A. Oliveras, and C. Tinelli. Solving SAT and SAT Modulo Theories: From an abstract Davis–Putnam–Logemann–Loveland procedure to DPLL (T). *Journal of the ACM*, 53(6):937–977, 2006.

J. Pearl. *Do We Need Higher-order Probabilities And, If So, what Do They Mean?* UCLA, Computer Science Department, 1987.

J. R. Shoenfield. *Mathematical logic*, volume 21. Addison-Wesley Reading, 1967.

J. Van Eijck and F. Schwarzentruber. Epistemic probability logic simplified. *Advances in Modal Logic*, 10:158–177, 2014.

Appendix A. Additional Proofs

Consistency is defined in the usual way: $\Delta \subseteq \mathsf{Glob}$ is *consistent* if $\Delta \nvdash_{(\Gamma,\Lambda)} \delta$ for some $\delta \in \mathsf{Glob}$. Note that a global formula of the form $\bigvee_{i=1}^{n} \delta_i$ is consistent if and only if there exists $1 \leq i \leq n$ such that δ_i is consistent.

Proof of Theorem 3.2 As usual, the proof of completeness follows by contraposition and consists on finding a model for the negation of an unprovable formula. Hence, we assume that $\nvdash_{(\Gamma,\Lambda)} \delta$ and build a F-structure satisfying $\neg\delta$. The construction combines several known techniques from equational logic, first-order logic and probabilistic logic, which interact in a non-trivial way. We begin by writing the consistent formula $\neg\delta$ in disjunctive normal form as $\psi_1 \vee \cdots \vee \psi_m$. Then, we choose a consistent disjunct ψ_j, of the form

$$\psi_j^1 \wedge \cdots \wedge \psi_j^{n_j}, \tag{2}$$

and define $\mathsf{RelF} = \{\psi_j^1, \ldots, \psi_j^{n_j}\} \subseteq \mathsf{Glob}$ to be the set of *relevant formulas* that should be satisfied in the final F-structure. We also add to the signature a new constant $c_{\varphi,n}$ for each $\varphi \in \mathsf{Loc}$ and $n \in N$, obtaining a signature $\mathrm{F}^+ = \{\mathrm{F}_n^+\}_{n\in\mathbb{N}}$ coinciding with F in all but $\mathrm{F}_0^+ = \mathrm{F}_0 \cup \left(\bigcup_{\varphi\in\mathsf{Loc}} \{c_{\varphi,n_0} \mid n_0 \in N\}\right)$.

Afterwards, we fix an enumeration for $\mathsf{Loc} \times \mathsf{Loc}$ and further extend the set RelF with witnesses for negated global formulas and with suitable certifications for non-negative global formulas, through the following inductive definition:

$$\begin{array}{lcl} W_0 & = & \mathsf{RelF} \\ W_{i+1} & = & W_i \cup \left\{ \neg\forall\varphi_i^1 \to \left(\forall[\neg\varphi_i^1]^{\tilde{n}}_{\tilde{c}_{\varphi_i^1}} \wedge \left(\forall\varphi_i^2 \to \forall[\varphi_i^2]^{\tilde{n}}_{\tilde{c}_{\varphi_i^1}} \right)\right)\right\} \quad \text{for each } i \in \mathbb{N}, \end{array}$$

where $\mathsf{names}(\varphi_i^1) \cup \mathsf{names}(\varphi_i^2) = \tilde{n} = \{n_1, \ldots, n_k\}$, $\tilde{c}_\varphi = \{c_{\varphi,n_1}, \ldots, c_{\varphi,n_k}\}$.

Lemma A.1. $W = \bigcup_{i\in\mathbb{N}} W_i$ is consistent (regarding F^+).

Proof of Lemma A.1 can be found in (Mor17) and follows the same steps as Henkin construction (Hen49).

We fix a maximal consistent extension Ξ of the set $W \subseteq \mathsf{Glob}^+$, whose existence is guaranteed by the Lindenbaum's Lemma. Then consider the F^+-algebra $\mathbb{A} = \mathbb{T}_{\mathrm{F}^+}(\varnothing)_{/\equiv}$, where the congruence relation is given by $t_1 \equiv t_2$ if $\forall(t_1 \approx t_2) \in \Xi$. The relation $\equiv$ is a congruence as consequence of axioms Eq1-4 and theorem N. A domain interpretation is then taken accordingly to the aforementioned maximal consistent set Ξ, $I^{\mathbb{A}}(D) = \{[t]_\equiv \mid \forall(t \in D) \in \Xi \text{ and } t \in T_{\mathrm{F}^+}(\varnothing)\}$ for each $D \in \mathcal{D}$.

- $\mathbb{A}$ **satisfies** Γ**:** by definition of $\equiv$, $\mathsf{E}(\Gamma)$, $\mathsf{C4}$, N, and recalling that Ξ is a maximal consistent set, it is easy to check that $\mathbb{A} \Vdash \Gamma$.
- $(\mathbb{A}, I^{\mathbb{A}})$ **verifies** Λ**:** given $\left(t_1 \in D_1, \ldots, t_{k_1} \in D_{k_1} \to t'_1 \circledcirc D'_1, \ldots, t'_{k_2} \circledcirc D_{k_2}\right) \in \Lambda$ and $\pi \in A^X$, notice that π results from applying a substitution $\sigma \in T_{\mathrm{F}^+}(\varnothing)^X$ and then a quotient by $\equiv$. Assume that $[\![t_i]\!]^\pi_{\mathbb{A}} \in I^{\mathbb{A}}(D_i)$ for each $1 \le i \le k_i$, which means that for each $1 \le i \le k_i$, $[\sigma(t_i)]_\equiv \in I^{\mathbb{A}}(D_i)$ or, equivalently, $\forall(\sigma(t_i) \in D_i) \in \Xi$. It means that $\forall(\sigma(t_1) \in D_1 \wedge \cdots \wedge \sigma(t_{k_1}) \in D_{k_1}) \in \Xi$, and from $\mathsf{E}(\Gamma)$ it follows that $\forall(\sigma(t'_1) \circledcirc D'_1 \vee \cdots \vee \sigma(t'_{k_2}) \circledcirc D'_{k_2}) \in \Xi$. But Ξ is maximal consistent with respect to the deductive system $\mathcal{H}_{(\Gamma,\Lambda)}$ and $\sigma(t'_1), \ldots, \sigma(t'_{k_2})$ are nameless terms, so it follows that exists $j \in \{1, \ldots, k_2\}$ such that $\forall(\sigma(t'_j) \circledcirc D'_j) \in \Xi$.

We note that each negated global formula in the maximal consistent set, $\neg\forall\varphi \in \Xi$, leads to an outcome $\rho^{\neg\forall\varphi} : N \to A$ assigning each name to the equivalence class of the appropriate constant: $\rho^{\neg\forall\varphi}(n) = [c_{\varphi,n}]_\equiv$. The set $S = \left\{\rho^{\neg\forall\varphi} \mid \neg\forall\varphi \in \Xi\right\}$ of possible outcomes is not empty since the conjugation of the reflexivity axiom $\mathsf{Eq1}$ with the axiom that enables the negation to be passed through the universal quantifier, $\mathsf{N2}$, implies that $\neg\forall(t \not\approx t) \in \Xi$, for each $t \in T(N)$.

A probability space is then defined, in the lines of (FHM90), and starts by choosing carefully a set of atoms of interest: initially we collect in Ω all the local formulas occurring inside probabilistic formulas of RelF, $\Omega = \bigcup_{\psi\in\mathsf{RelF}\cap(\mathsf{Prob}\cup\neg\mathsf{Prob})} \mathsf{lnPr}(\psi)$, where $\mathsf{lnPr}(q_1{\cdot}\mathsf{Pr}(\varphi_1){+}\ldots{+}q_l{\cdot}\mathsf{Pr}(\varphi_l) \ge b) = \mathsf{lnPr}(\neg(q_1{\cdot}\mathsf{Pr}(\varphi_1){+}\ldots{+}q_l{\cdot}\mathsf{Pr}(\varphi_l) \ge b)) = \{\varphi_1, \ldots, \varphi_l\}$, and then use it to define the suitable atoms $\Theta = \left\{ \bigwedge_{\gamma\in\Upsilon} \gamma \wedge \bigwedge_{\omega\in\Omega\setminus\Upsilon} \neg\omega \mid \Upsilon \subseteq \Omega \right\}$. We consider a representative outcome for each element of $\theta \in \Theta$, whenever it is possible: if $S^\theta \neq \varnothing$, choose $\rho_\theta \in S^\theta$ and let us represent the probability assigned to ρ_θ by x_θ; otherwise, if $S^\theta = \varnothing$, i.e. $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \forall\neg\theta$, fix $x_\theta = 0$. The accuracy of Θ immediately implies that $\bigcup_{\theta\in\Theta} S^\theta = S$ and $S^{\theta_1} \cap S^{\theta_2} = \varnothing$, for each $\theta_1 \neq \theta_2$. The set Θ has the crucial local formulas to define the system of inequalities

$$\begin{cases} a_1 x_{\varphi_1} + \cdots + a_k x_{\varphi_k} \geq b, & \text{for each } a_1 \mathsf{Pr}(\varphi_1) + \cdots + a_k \mathsf{Pr}(\varphi_k) \geq b \in \mathsf{RelF} \\ a_1 x_{\varphi_1} + \cdots + a_k x_{\varphi_k} < b, & \text{for each } a_1 \mathsf{Pr}(\varphi_1) + \cdots + a_k \mathsf{Pr}(\varphi_k) < b \in \mathsf{RelF} \\ \sum_{\theta \in \Theta \text{ st } \theta \to \varphi} x_\theta = x_\varphi, & \text{for each } \varphi \in \Omega \\ \sum_{\theta \in \Theta} x_\theta = 1 & \\ x_\theta = 0, & \text{for each } \theta \in \Theta \text{ such that } S^\theta = \varnothing \\ x_\theta \geq 0, & \text{for all } \theta \in \Theta \end{cases} \tag{3}$$

We claim that this system of inequalities has a solution. Indeed, using Fagin, Halpern and Megiddo's result of soundness and completeness for the axioms of inequality (see Section 4 of (FHM90)), we know that (3) is unsatisfiable if and only if it is inconsistent. But it leads to a contradiction, as we can find a global formula that represents this system of inequalities within DEqPrL. Let us look at this in more detail.

To write down a global formula that represents the system (3), let us fix an order on elements of Ω: $\Omega = \{\varphi_1, \ldots, \varphi_{|\Omega|}\}$. Then, consider the successive application of axiom P2 to deduce that

$$\begin{aligned} \mathsf{Pr}(\varphi_1) &= \mathsf{Pr}(\varphi_1 \wedge \varphi_2) + \mathsf{Pr}(\varphi_1 \wedge \neg\varphi_2) = \\ &= \mathsf{Pr}(\varphi_1 \wedge \varphi_2 \wedge \varphi_3) + \mathsf{Pr}(\varphi_1 \wedge \varphi_2 \wedge \neg\varphi_3) + \mathsf{Pr}(\varphi_1 \wedge \neg\varphi_2 \wedge \varphi_3) + \mathsf{Pr}(\varphi_1 \wedge \neg\varphi_2 \wedge \neg\varphi_3) = \\ &= \ldots = \sum_{\theta \in \Theta \text{ st } \theta \to \varphi_1} \mathsf{Pr}(\theta). \end{aligned} \tag{4}$$

It means that $\vdash_{(\Gamma,\Lambda)} \mathsf{Pr}(\varphi_1) = \sum_{\theta \in \Theta \text{ st } \theta \to \varphi_1} \mathsf{Pr}(\theta)$. We can obtain a similar formula for each $\varphi \in \Omega$. Moreover, since $\bigvee_{\theta \in \Theta} \theta \leftrightarrow \top$ and $\theta_i \wedge \theta_j \leftrightarrow \bot$ for any $\theta_i, \theta_j \in \Theta$, $\theta_i \neq \theta_j$, using axioms P2 and P4 we can deduce that $\mathsf{Pr}\left(\bigvee_{\theta \in \Theta} \theta\right) = \sum_{\theta \in \Theta} \mathsf{Pr}(\theta)$ and it follows that $\sum_{\theta \in \Theta} \mathsf{Pr}(\theta) = 1$. Before finishing, notice that PAux2 and P2 imply that: $\vdash_{(\Gamma,\Lambda)} \bigwedge_{\theta \in \Theta} (\forall(\neg\theta) \to \mathsf{Pr}(\theta) = 0)$.

Axiom P1 and the previous justifications, allow us to write (2) equivalently as:

$$\psi_j^1 \wedge \cdots \wedge \psi_j^{n_j} \wedge \bigwedge_{\varphi \in \Omega} \left(\mathsf{Pr}(\varphi) = \sum_{\substack{\theta \in \Theta \\ \theta \to \varphi}} \mathsf{Pr}(\theta) \right) \wedge \left(\sum_{\theta \in \Theta} \mathsf{Pr}(\theta) = 1 \right) \wedge \left(\bigwedge_{\theta \in \Theta} \forall(\neg\theta) \to \mathsf{Pr}(\theta) = 0 \right) \wedge \bigwedge_{\theta \in \Theta} (\mathsf{Pr}(\theta) \geq 0). \tag{5}$$

Since we can assign probabilities independently to the different elements in Θ, (5) is satisfiable if and only if the system of inequalities (3) is satisfiable. Under the hypothesis that the system of inequalities is unsatisfiable, using the results of soundness and completeness for the axioms of inequality, the system would be inconsistent. But it would mean that we could derive an inconsistency from (5) using I1-I6, C1-C4, which is a contradiction with the consistency of (2). We conclude that the system (3) is satisfiable. Let $\{x_\theta^*\}_{\theta \in \Theta}$ be a solution. The solution of (3) is used to define a probability distribution over the atoms and thus over the outcomes satisfying them. The probability distribution $\mathcal{P} : \mathsf{S} \to [0, 1]$ is defined as follows:

$$\begin{cases} \mathcal{P}(\rho_\theta) = x_\theta^*, & \text{for each } \theta \in \Theta, \\ \mathcal{P}(\rho) = 0, & \text{for each } \rho \in \mathsf{S} \setminus \{\rho_\theta \mid \theta \in \Theta\}. \end{cases}$$

A probability space $\mathbb{P} = (S, \mathscr{A}, \mu)$ is built on top of this probability distribution, considering the σ-algebra $\mathscr{A}$ generated by the set $\{S^\varphi \mid \varphi \in \mathsf{Loc}\}$ and the probability measure $\mu : \mathscr{A} \to [0,1]$ such that $\mu(X) = \sum_{\rho\in X} \mathcal{P}(\rho)$. Let us verify that μ is a probability measure:

- Given $X \in \mathscr{A}$, $\mu(X) \geq 0$ since $\mu(X) = \sum_{\rho\in X}\mathcal{P}(\rho)$, and the system of inequalities (3) together with the definition of $\mathcal{P}$ imply that $\mathcal{P}(\rho) \geq 0$ for each $\rho \in S$.
- We conclude that $\mu(S) = 1$ by observing that $S \in \mathscr{A}$ as result of $S = S^{t\approx t}$, and further $\mu(S) = \sum_{\rho\in S} \mathcal{P}(\rho)$, which leads to the expected measure 1 for the entire set of possible outcomes by simply recalling the definition of $\mathcal{P}$ and writing $\mu(S) = \sum_{\rho\in S}\mathcal{P}(\rho) = \sum_{\rho\in S\setminus\{\rho^\theta|\theta\in\Theta\}}\mathcal{P}(\rho) + \sum_{\theta\in\Theta}\mathcal{P}(\rho_\theta) = 0 + \sum_{\theta\in\Theta} x^*_\theta$. Since $\{x^*_\theta\}_{\theta\in\Theta}$ is a solution for (3) we actually have $\mu(S) = \sum_{\theta\in\Theta} x^*_\theta = 1$.
- Given a countable collection of pairwise disjoint sets $\{X_i\}_{i\in I} \subseteq \mathscr{A}$, the equality $\mu\left(\bigcup_{i\in I} X_i\right) = \sum_{i\in I}\mu(X_i)$ holds as a consequence of sets $\{X_i\}_{i\in I}$ being pairwise disjoint and from the following equalities: $\sum_{i\in I}\mu(X_i) = \sum_{i\in I}\sum_{\rho\in X_i}\mathcal{P}(\rho) = \sum_{\rho\in\bigcup_{i\in I}X_i}\mathcal{P}(\rho) = \mu\left(\bigcup_{i\in I}X_i\right)$.

  Just note that each of the previous sums have a finite number of non-zero elements.

Now that a F-structure $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P})$ has emerged, it remains to prove that it actually satisfies all the relevant formulas RelF. For that purpose, we leave an auxiliary remark whose proof follows easily by induction on the complexity of φ.

Remark 1. Given $\neg\forall\varphi_0 \in \Xi$ and a local formula $\varphi \in \mathsf{Loc}$ with $\mathsf{names}(\varphi) = \tilde{n}$,

$$\forall[\varphi]^{\tilde{n}}_{\tilde{c}_{\varphi_0}} \in \Xi \text{ if and only if } \mathbb{A}, I^{\mathbb{A}} \Vdash [\varphi]^{\tilde{n}}_{\tilde{c}_{\varphi_0}}.$$

We conclude the proof verifying that we have indeed a model for RelF. Recall that $\mathsf{RelF} \subseteq \forall\mathsf{Loc} \cup \neg\forall\mathsf{Loc} \cup \mathsf{Prob} \cup \neg\mathsf{Prob}$, consider $\gamma \in \mathsf{RelF}$ and let us analyze the four cases:

- if γ is of the form $\forall\varphi$ with $\mathsf{names}(\varphi) = \tilde{n}$, we want to prove that for every $\rho \in S$, $(\mathbb{A}, I^{\mathbb{A}}), \rho \Vdash_{\mathsf{loc}} \varphi$. Given $\rho \in S$, recall that it was motivated by some $\neg\forall\varphi_0 \in \Xi$, say that $\rho = \rho^{\neg\forall\varphi_0}$. Since $\forall\varphi \in \mathsf{RelF} \subseteq \Xi$ it follows that $\forall[\varphi]^{\tilde{n}}_{\tilde{c}_{\varphi_0}} \in \Xi$ by construction of W. Using Remark 1 we conclude that $\mathbb{A}, I^{\mathbb{A}} \Vdash [\varphi]^{\tilde{n}}_{\tilde{c}_{\varphi_0}}$, which according to definition of $\rho^{\neg\forall\varphi_0}$ implies that $(\mathbb{A}, I^{\mathbb{A}}), \rho^{\neg\forall\varphi_0} \Vdash_{\mathsf{loc}} \varphi$.
- if γ is of the form $\neg\forall\varphi$, with $\mathsf{names}(\neg\varphi) = \mathsf{names}(\varphi) = \tilde{n}$, notice that $\rho^{\neg\forall\varphi} \in S$. Moreover, since $\neg\forall\varphi \in \Xi$, it follows that $\forall[\neg\varphi]^{\tilde{n}}_{\tilde{c}_\varphi} \in \Xi$. Remark 1 implies that $\mathbb{A}, I^{\mathbb{A}} \Vdash [\neg\varphi]^{\tilde{n}}_{\tilde{c}_\varphi}$, which by definition of $\rho^{\neg\forall\varphi}$ leads to $(\mathbb{A}, I^{\mathbb{A}}), \rho^{\neg\forall\varphi} \Vdash_{\mathsf{loc}} \neg\varphi$, so $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \neg\forall\varphi$.
- If $\gamma \in \mathsf{Prob}$ is of the form $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_l \cdot \mathsf{Pr}(\varphi_l) \geq b$, we have:

$$\begin{aligned}
& (\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_l \cdot \mathsf{Pr}(\varphi_l) \geq b \\
\text{iff}\quad & q_1 \cdot \mu(S^{\varphi_1}) + \cdots + q_l \cdot \mu(S^{\varphi_l}) \geq b \\
\text{iff}\quad & q_1 \sum_{\substack{\rho\in S \text{ st} \\ (\mathbb{A}, I^{\mathbb{A}}),\rho\Vdash_{\mathsf{loc}}\varphi_1}} \mathcal{P}(\rho) + \cdots + q_l \sum_{\substack{\rho\in S \text{ st} \\ (\mathbb{A}, I^{\mathbb{A}}),\rho\Vdash_{\mathsf{loc}}\varphi_l}} \mathcal{P}(\rho) \geq b \\
\text{iff}\quad & q_1 \sum_{\substack{\theta\in\Theta \text{ st} \\ \theta\to\varphi_1}} \mathcal{P}(\rho_\theta) + \cdots + q_l \sum_{\substack{\theta\in\Theta \text{ st} \\ \theta\to\varphi_l}} \mathcal{P}(\rho_\theta) \geq b \\
\text{iff}\quad & q_1 \sum_{\substack{\theta\in\Theta \text{ st} \\ \theta\to\varphi_1}} x^*_\theta + \cdots + q_l \sum_{\substack{\theta\in\Theta \text{ st} \\ \theta\to\varphi_l}} x^*_\theta \geq b.
\end{aligned}$$

The last inequality is valid since $q_1 \cdot \mathsf{Pr}(\varphi_1) + \cdots + q_l \cdot \mathsf{Pr}(\varphi_l) \geq b \in \mathsf{RelF}$ and $\{x^*_\theta\}_{\theta \in \Theta}$ is a solution for (3), hence the first assertion holds as well.

- If $\gamma \in \neg\mathsf{Prob}$, the reasoning is similar.

Hence we have: $(\mathbb{A}, I^{\mathbb{A}}, \mathbb{P}) \Vdash \neg\delta$. □

Sketch of the proof of Lemma 4.1 For lack of space, we only summarize the idea underlying the proof of Lemma 4.1 very briefly. The details can be found in (Mor17).

To prove the correctness of Algorithm 1, one should ensure that every model in the equational context corresponds to a valuation over the wider propositional set of variables $\mathcal{B}^*$, and vice-versa. In this sense, for the direct implication, we construct several valuations from outcomes in a model for δ and then unify them in a bigger valuation. The verification that the Algorithm returns Sat is an immediate consequence of the construction. For the reciprocal implication, we split the bigger valuation into valuations over $\widetilde{\mathcal{B}}$, and then over $\mathcal{B}$. Finally, we use an argument similar to the one used for the proof of completeness to construct the final model for δ. □