EPiC Series in Computing

Volume 50, 2017, Pages 188–201

GCAI 2017. 3rd Global Conference on Artificial Intelligence

EPiC Computing

Secrecy-Preserving Reasoning and Query Answering in Probabilistic Description Logic $Prob\text{-}\mathcal{EL}^{>0,=1}_{-\top}$ KBs

Gopalakrishnan Krishnasamy-Sivaprakasam[1], Adrienne Raglin[2], Douglas Summers-Stay[2], and Giora Slutzki[3]

[1] Central State University, Wilberforce, Ohio, USA
gkrishnasamy@centralstate.edu

[2] U.S. Army Research Laboratory, Adelphi, Maryland, USA
{adriene.j.raglin.civ; douglas.a.summers-stay.civ}@mail.mil

[3] Iowa State University, Ames, Iowa, USA
slutzki@iastate.edu

Abstract

In this paper we study Secrecy-Preserving Query Answering problem under the Open World Assumption (OWA) for $Prob\text{-}\mathcal{EL}^{>0,=1}_{-\top}$ Knowledge Bases (KBs). We have designed a tableau procedure to compute a *semi-model* $\mathbb{M}$ *over* the given KB which eventually is equivalent to a probabilistic model to KB. A semi-model over KB is a tuple $\mathbb{M} = \langle \Omega, \mathbb{W}, \tau \rangle$ where Ω is a finite set $\{\omega_0, \omega_1, ..., \omega_m\}$, $\mathbb{W}$ is a function that assigns each element in Ω with a set of assertions whose concepts occur in KB and τ is a function that assigns each element in Ω with a real number such that $\tau(\omega_0) = 0$ and $\tau(\omega_i) > 0$ for all $i \in \Omega \setminus \{0\}$. Given a secrecy set $\mathbb{S}$, which is a finite set of assertions, we compute a function E, called an envelope of $\mathbb{S}$, which assigns a set $E(\omega)$ of assertions to each $\omega \in \Omega$. E provides logical protection to the secrecy set $\mathbb{S}$ against the reasoning of a querying agent. Once the semi-model $\mathbb{M}$ and an envelope E are computed, we define the secrecy-preserving semi-model $\mathbb{M}_E$. Based on the information available in $\mathbb{M}_E$, assertional queries with probabilistic operators can be answered efficiently while preserving secrecy. To the best of our knowledge, this work is first one studying secrecy-preserving reasoning in description logic augmented with probabilistic operators. When the querying agent asks a query q, the reasoner answers "Yes" if information about q is available in $\mathbb{M}_E$; otherwise, the reasoner answers "Unknown". Being able to answer "Unknown" plays a key role in protecting secrecy under OWA. Since we are not computing all the consequences of the knowledge base, answers to the queries based on just secrecy-preserving semi-model $\mathbb{M}_E$ could be erroneous. To fix this problem, we further augment our algorithms by providing recursive query decomposition algorithm to make the query answering procedure foolproof.

1 Introduction

In web based business activities, a major issue is how to protect private information of users from the unauthorized users while making sure the smooth transfer of non confidential information that available in the public domain among all the users. In the literature, most of the approaches dealing with "information protection" are based on access control mechanisms. Controlled query evaluation (CQE)

C. Benzmüller, C. Lisetti and M. Theobald (eds.), GCAI 2017 (EPiC Series in Computing, vol. 50), pp. 188–201

is an approach to enforcing secrecy based on control access mechanisms. Biskup et al. in [2, 1] studied the problem of enforcing secrecy using CQE on complete relational databases. Further, in [4, 3], authors extended their research to incomplete databases. Since description logics (DLs) underlie web ontology languages (OWLs), recently Cuenca Grau et.al. in [6, 7] adapted the CQE approach to study secrecy-preserving reasoning in DL knowledge bases (KBs) which are assumed to be incomplete.

In [15, 17], the authors have developed a secrecy framework that attempts to satisfy the following competing goals: (a) it protects secret information and (b) queries are answered as informatively as possible (subject to satisfying property (a)). The notion of an *envelope* to hide secret information against logical inference was first defined and used in [15]. In [17], Tao et al., introduced a more elaborate conceptual framework for secrecy-preserving query answering (SPQA) under Open World Assumption (OWA) with multiple querying agents. This approach is based on OWA and (so far) it has been restricted to instance-checking queries. Specifically, in [15, 17] the main idea was to utilize the secret information within the reasoning process, but then answering "Unknown" whenever the answer is truly unknown or in case the true answer could compromise confidentiality. In [11, 13] we have extended the work of Tao et al., reported in [15], to the $\mathcal{ELH}$ and $\mathcal{EL}^+$ languages respectively, and studied secrecy in the context of assertions as well as general concept inclusions (GCIs). Finally, in [14] we have studied the SPQA problem in modalized DL KBs.

The Prob-DLs are DLs with probabilistic constructors. Lutz et al., in [12] introduced a new family of probabilistic DLs based on probabilistic FOL which is discussed in [9]. In [12], the authors argued that Prob-DLs are suitable logical formalism to model bio medical domain knowledge. Further, in [8] Gutierrez-Basulto et.al., studied various reasoning tasks in *Prob-*$\mathcal{EL}$. Secrecy of probabilistic information has been studied by many researchers in different contexts. Discussions of privacy issues in probabilistic knowledge models were reported in [5, 10]. Motivated by these works, in this paper we study SPQA problem in Prob-DL *Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$ KBs. *Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$ is a top-free description logic $\mathcal{EL}$ augmented with the probabilistic operators $P_{>0}$ and $P_{=1}$. The reason for excluding $\top$ from the syntax of *Prob-EL*$^{>0,=1}$ logic is to avoid computing tautological statements that are not relevant to secrecy preservation. In the literature there are several top-free DL languages, for instance, *DL-Lite*$_R$ is a top-free DL, see (Calvanese et al., 2007). The syntax and semantics of the *Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$ are presented in Section 2.

As a first step in constructing SPQA system, we design a tableau algorithm to compute a finite collection of sets of assertions referred to as *semi-model*. One of the sets in this collection contains a set of consequences of the given KB $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$, restricted to concepts that actually occur in Σ and an extra "auxiliary" set of concepts defined over the signature of Σ. This semi-model, once computed, remains fixed and is not modified. The tableau algorithm is sound and complete under the restrictions stated above, see Section 3. Since the computed semi-model does not contain all the consequences of the KB, in order to answer user queries we have designed recursive algorithms which break the queries into smaller assertions all the way until the information in the semi-model can be used. In effect, we have split the task of query answering into two parts: in the first part we compute all the consequences of Σ restricted to concepts and individuals that occur in Σ, in the second part we use a recursive algorithm to evaluate more complex queries with the base case that has been computed in the first part.

To protect the secret information in the secrecy set $\mathbb{S}$, we compute an envelope E which is a function that defines each set in the computed semi-model a set of assertions. This envelope is computed by a tableau algorithm based on the idea of inverting the local and global expansion rules given in Figures 1 and 2. The idea behind the envelope concept is that no expression in the envelope can be logically deduced from information outside the envelope. Once such envelope is computed, the answers to the queries are censored whenever the queries belong to the envelope. Since, generally, an envelope for a given secrecy set is not unique, the developer can force the algorithm to output a specific envelope from the available choices satisfying the needs of application domain, company policy, social obligations and user preferences.

Next, we discuss query answering procedures which allow us answer queries without revealing secrets. The queries are answered based on the information available in the secrecy-preserving semi-model, see Section 4. Usually in SPQA framework queries are answered by checking their membership in the computed semi-model. Since the secrecy-preserving semi-model does not contain all the statements entailed by Σ, we need to extend the query answering procedure from just membership checking. Towards that end we have designed a recursive algorithm to answer more complicated queries. To answer a query q, the algorithm first checks if q is a member secrecy-preserving semi-model (in a particular set that contains the consequences of Σ), in which case the answer is "Yes"; otherwise, the given query is broken into subqueries based on the constructors, and the algorithm is applied recursively on the subqueries, see Section 5.

2 Syntax and Semantics of *Prob*-$\mathcal{EL}^{>0,=1}_{-\top}$

A vocabulary of $\mathcal{EL}$ is a triple $< N_O, N_C, N_R >$ of countably infinite, pair-wise disjoint sets. The elements of N_O are called *objects* or *individuals*, the elements of N_C are called *concept names* and the elements of N_R are called *role names*. The set of $\mathcal{EL}$ *concepts* is denoted by $\mathcal{C}$ and it is syntactically defined by the following rules

$$C ::= A \mid \top \mid C \sqcap D \mid \exists r.C$$

where $A \in N_C$, $r \in N_R$, $\top$ denotes the "*top concept*", and $C, D \in \mathcal{C}$. The semantics of $\mathcal{EL}$ concepts is specified, as usual, by an *interpretation* $\mathcal{I} = (\Delta, \cdot^{\mathcal{I}})$ where Δ is a non-empty *domain* of the interpretation, and $\cdot^{\mathcal{I}}$ is an *interpretation function* mapping each $a \in N_O$ to an element $a^{\mathcal{I}} \in \Delta$, each $A \in N_C$ to a subset $A^{\mathcal{I}} \subseteq \Delta$, and each $r \in N_R$ to a binary relation $r^{\mathcal{I}} \subseteq \Delta \times \Delta$. The interpretation function $\cdot^{\mathcal{I}}$ is extended inductively to all $\mathcal{EL}$ concepts in the usual manner:

$$\begin{aligned} \top^{\mathcal{I}} &= \Delta \\ (C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\ (\exists r.C)^{\mathcal{I}} &= \{d \in \Delta \mid \exists e \in C^{\mathcal{I}} : (d,e) \in r^{\mathcal{I}}\} \end{aligned}$$

In [12], the authors (did not study secrecy at all) did show that some extensions of DL $\mathcal{EL}$ with probabilistic constructors are intractable. In this paper we initiate the study of secrecy-preserving reasoning in probabilistic KBs, specifically in the language *Prob*-$\mathcal{EL}^{>0,=1}_{-\top}$ which is a fragment of the language Prob-$\mathcal{EL}^{01}_{c}$ introduced in [12]. The set of *Prob*-$\mathcal{EL}^{>0,=1}_{-\top}$ concepts is denoted by $\mathcal{PC}$, and the concepts are formed according to the following syntax

$$C ::= A \mid C \sqcap D \mid \exists r.C \mid P_{>0}C \mid P_{=1}C$$

where $C, D \in \mathcal{PC}$. *Assertions* are expressions of the form $C(a)$ or $r(a,b)$ and *general concept inclusions (GCIs)* are expressions of the form $C \sqsubseteq D$ where $C, D \in \mathcal{PC}$, $r \in N_R$ and $a, b \in N_O$.

To define the semantics of *Prob*-$\mathcal{EL}^{>0,=1}_{-\top}$, we use probabilistic interpretations which combine the $\mathcal{EL}$ interpretations together with an extra probabilistic structure, see [12, 8]. Formally, a *probabilistic interpretation* is a structure $\mathcal{I} = (\Delta, W, \{\mathcal{I}_\omega\}_{\omega \in W}, \mu)$ where Δ is a non-empty domain, W a non-empty set of possible worlds, μ a discrete probability measure on W, and for each $\omega \in W$, $\mathcal{I}_\omega$ is a classical $\mathcal{EL}$ interpretation. Here we assume that for each world ω in W, $\mathcal{I}_\omega$ have the same domain Δ. We shall write $C^{\mathcal{I},\omega}$ or even C^{ω} (respectively $r^{\mathcal{I},\omega}$ or r^{ω}) for $C^{\mathcal{I}_\omega}$ (respectively $r^{\mathcal{I}_\omega}$). We make the *global name assumption (gna)*, namely, that individual names are interpreted *globally*, in a world-independent fashion: for all $a \in N_O$ and all $u, v \in W$, $a^{\mathcal{I},u} = a^{\mathcal{I},v}$. Define the probability of a domain element to belong to a concept name by $p_d^{\mathcal{I}}(A) = \mu(\{\omega \in W \mid d \in A^{\mathcal{I},\omega}\})$ and extend it to compound concepts and simultaneously, in a mutually recursive fashion, define the extension of the interpretation function $\cdot^{\mathcal{I},\omega}$ to compound concepts as follows:

$$\begin{aligned} p_d^{\mathcal{I}}(C) &= \mu(\{\omega \in W \mid d \in C^{\mathcal{I},\omega}\}), \text{ for all compound concepts } C, \text{ and}\\ (C \sqcap D)^{\mathcal{I},\omega} &= C^{\mathcal{I},\omega} \cap D^{\mathcal{I},\omega},\\ (\exists r.C)^{\mathcal{I},\omega} &= \{d \in \Delta \mid \exists e \in C^{\mathcal{I},\omega} : (d,e) \in r^{\mathcal{I},\omega}\},\\ (P_{>0}C)^{\mathcal{I},\omega} &= \{d \in \Delta \mid p_d^{\mathcal{I}}(C) > 0\}, \text{ and}\\ (P_{=1}C)^{\mathcal{I},\omega} &= \{d \in \Delta \mid p_d^{\mathcal{I}}(C) = 1\}. \end{aligned}$$

An *ABox* $\mathcal{A}$ is a finite, non-empty set of assertions and a *TBox* $\mathcal{T}$ is a finite set of GCIs. A *Prob-*$\mathcal{EL}_{-\top}^{>0,=1}$ KB is a pair $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ where $\mathcal{A}$ is an ABox and $\mathcal{T}$ is a TBox. Let $\mathcal{I} = (\Delta, W, \{\mathcal{I}_\omega\}_{\omega \in W}, \mu)$ be a probabilistic interpretation, $\omega \in W$, $C, D \in \mathcal{PC}$, $r \in N_R$ and $a, b \in N_O$. We say that $(\mathcal{I}, \omega)$ *satisfies* $C(a)$, $r(a,b)$, or $C \sqsubseteq D$, notation $(\mathcal{I},\omega) \models C(a)$, $(\mathcal{I},\omega) \models r(a,b)$ or $(\mathcal{I},\omega) \models C \sqsubseteq D$ if, respectively, $a^{\mathcal{I},\omega} \in C^{\mathcal{I},\omega}$, $(a^{\mathcal{I},\omega}, b^{\mathcal{I},\omega}) \in r^{\mathcal{I},\omega}$ or $C^{\mathcal{I},\omega} \subseteq D^{\mathcal{I},\omega}$. $(\mathcal{I},\omega)$ *satisfies* Σ, notation $(\mathcal{I},\omega) \models \Sigma$, if $(\mathcal{I},\omega)$ satisfies all the assertions in $\mathcal{A}$ and all the GCIs in $\mathcal{T}$. $\mathcal{I}$ *satisfies* Σ, or $\mathcal{I}$ is a *model* of Σ, if there exists a $\omega \in W$ such that $(\mathcal{I},\omega) \models \mathcal{A}$ and for all $\omega \in W$, $(\mathcal{I},\omega) \models \mathcal{T}$, see [12].

Definition 1. *Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be an Prob-$\mathcal{EL}_{-\top}^{>0,=1}$ KB and let α be an assertion. We say that Σ entails α, notation $\Sigma \models \alpha$, if for all probabilistic interpretations $\mathcal{I} = (\Delta, W, \{\mathcal{I}_\omega\}_{\omega \in W}, \mu)$ satisfying Σ, $(\mathcal{I},\omega) \models \Sigma \Rightarrow (\mathcal{I},\omega) \models \alpha$, for all $\omega \in W$ of $\mathcal{I}$.*

Note that by Definition 1, any probabilistic interpretation $\mathcal{I}$ that satisfies Σ must satisfy the TBox $\mathcal{T}$ in every world $\omega \in W$. On the other hand. the ABox $\mathcal{A}$ must be satisfied in some world $\omega \in W$. In other words, a probabilistic interpretation $\mathcal{I}$ satisfies KB Σ in some world in $\omega \in W$. For more details see [12]. In section 3, we have designed a tableau algorithm to compute a model for the Σ which has the above mentioned characteristics.

3 Computation of a model for *Prob*-$\mathcal{EL}_{-\top}^{>0,=1}$ KB

Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be an *Prob-*$\mathcal{EL}_{-\top}^{>0,=1}$ KB. Denote by N_Σ the set of all concept names and role names occurring in Σ and let $\mathcal{S}$ be a finite set of concepts over the symbol set N_Σ[1]. Let $\mathcal{C}_{\Sigma,\mathcal{S}}$ be the set of all subconcepts of concepts that occur in $\mathcal{S}$ or Σ and define

$$\mathcal{A}^* = \{C(a) \mid C \in \mathcal{C}_{\Sigma,\mathcal{S}} \text{ and } \Sigma \models C(a)\} \cup \{r(a,b) \mid \Sigma \models r(a,b)\}.$$

We use $\mathcal{O}_\Sigma$ to denote the set of individual names that occur in Σ, and define the *witness set* $\mathcal{W} = \{w_C^r \mid r$ is a role name that occurs in Σ and $C \in \mathcal{C}_{\Sigma,\mathcal{S}}\}$. Define $\mathcal{O}^* = \mathcal{O}_\Sigma \cup \mathcal{W}$. Typically, in studying reasoning problems, the goal of designing a reasoning algorithm is to compute a model for given KB. For a given probabilistic KB Σ, we design a tableau algorithm that builds a *semi-model over Σ* which eventually is equivalent to a probabilistic model to Σ. A semi-model over Σ is a tuple $\mathbb{M} = \langle \Omega, \mathbb{W}, \tau \rangle$ where Ω is a finite set $\{\omega_0, \omega_1, ..., \omega_m\}$, $\mathbb{W}$ is a function that assigns each element in Ω with a set of assertions whose concepts occur in Σ and τ is a function that assigns each element in Ω with a real number such that $\tau(\omega_0) = 0$ and $\tau(\omega_i) > 0$ for all $i \in \Omega \setminus \{0\}$.

Given Σ and $\mathcal{C}_{\Sigma,\mathcal{S}}$, we outline a procedure that computes a semi-model over Σ, see [12] for a similar construction. We use $\mathbb{P}^0$ to denote the set of subconcepts of the form $P_{>0}C$ of concepts that occur in $\mathcal{C}_{\Sigma,\mathcal{S}}$.

[1] A technicality: $\mathcal{S}$ will be used in Section 4 in the context of secrecy-preserving reasoning.

$\sqcap^+ -$rule : if $C(a),\ D(a) \in \mathbb{W}(\omega),\ C \sqcap D \in \mathcal{C}_{\Sigma,\mathcal{S}}$, and
$C \sqcap D(a)\ \notin \mathbb{W}(\omega)$, then $\mathbb{W}(\omega) := \mathbb{W}(\omega) \cup \{C \sqcap D(a)\}$;
$\sqcap^- -$rule : if $C \sqcap D(a) \in \mathbb{W}(\omega)$, and $C(a) \notin \mathbb{W}(\omega)$ or $D(a)\ \notin \mathbb{W}(\omega)$,
then $\mathbb{W}(\omega) := \mathbb{W}(\omega) \cup \{C(a), D(a)\}$;
$\exists^+ -$rule : if $r(a,b),\ C(b) \in \mathbb{W}(\omega),\ \exists r.C \in \mathcal{C}_{\Sigma,\mathcal{S}}$ and
$\exists r.C(a)\ \notin \mathbb{W}(\omega)$, then $\mathbb{W}(\omega) := \mathbb{W}(\omega) \cup \{\exists r.C(a)\}$;
$\exists^- -$rule : if $\exists r.C(a) \in \mathbb{W}(\omega)$, and $\forall b \in \mathcal{O}^*, \{r(a,b), C(b)\} \not\subseteq \mathbb{W}(\omega)$,
then $\mathbb{W}(\omega) := \mathbb{W}(\omega) \cup \{r(a,w_C^r), C(w_C^r)\}$, where $w_C^r \in \mathcal{W}$;
$\sqsubseteq\ -$rule : if $C(a) \in \mathbb{W}(\omega),\ C \sqsubseteq D \in \mathcal{T}$, and $D(a)\ \notin \mathbb{W}(\omega)$,
then $\mathbb{W}(\omega) := \mathbb{W}(\omega) \cup \{D(a)\}$.

Figure 1: Local expansion rules.

Let $\Omega = \{0\} \cup \mathbb{P}^0$, and $\tau(0) = 0$ and $\tau(\omega) = 1/|\Omega \setminus \{0\}|$ for all $\omega \in \Omega \setminus \{0\}$. The procedure starts computing $\mathbb{M}$ with the initialization step: The set $\mathbb{W}(0)$ is initialized as the ABox $\mathcal{A}$, and $\mathbb{W}(\omega)$ is initialized as $\emptyset$ for all $\omega \in \Omega \setminus \{0\}$. Further, $\mathbb{M}$ is computed by recursively applying the *expansion rules* in Figures 1 and 2. $\mathbb{M}$ is said to be *completed* if no expansion rule in Figures 1 or 2 is applicable to it. The procedure is designed to output a completed semi-model $\mathbb{M}$ with $\mathbb{W}(0) = \mathcal{A}^*$. The strategies for designing this procedure are (a) the set $\mathbb{W}(0)$ collects the consequences of the given KB and (b) the members of the sets $\mathbb{W}(\omega)$ where $\omega \in \Omega \setminus \{0\}$ serve as witnesses for the some of the members in the set $\mathbb{W}(0)$. For the purpose of query answering, $\mathbb{M}$ is used as a "good approximation" of a canonical model of the given KB, see Section 5.

In more detail, there are two kinds of expansion rules: (a) *local* expansion rules and (b) *global* expansion rules. Local expansion rules are given in Figure 1 and generate new assertions within a particular set $\mathbb{W}(\omega)$. The $\sqcap^-$-rule decomposes conjunctions, and $\exists^-$-rule decomposes existential restriction assertions of the form $\exists r.C(a)$ by introducing a corresponding witness w_C^r from the set $\mathcal{W}$. The $\sqsubseteq$-rule derives new assertions based on the GCIs present in $\mathcal{T}$. Finally, to construct concept assertions whose associated concept expressions already belong to $\mathcal{C}_{\Sigma,\mathcal{S}}$, we use the $\sqcap^+$ and $\exists^+$-rules. The global expansion rules are given in Figure 2. The $P_{>0}^+$ and $P_{=1}^+$-rules add new probabilistic assertions in all the sets $\mathbb{W}(\nu)$ for $\nu \in \Omega$ if the corresponding probabilistic concept expressions already occur in $\mathcal{C}_{\Sigma,\mathcal{S}}$. The $P_{>0}^-$-rule generates an assertion to the set $\mathbb{W}(P_{>0}C)$ if $P_{>0}C(a)$ occurs in the set $\mathbb{W}(\omega)$ for some $\omega \in \Omega$. Similarly, the $P_{=1}^-$-rule generates an assertion to the each set $\mathbb{W}(w)$ for $\omega \in \Omega \setminus \{0\}$ if $P_{=1}C(a)$ occurs in the set $\mathbb{W}(\omega)$ for some $\omega \in \Omega$.

Example 1. *Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be a Prob-$\mathcal{EL}_{-\top}^{>0,=1}$ KB, where $\mathcal{A} = \{P_{>0}A(a),\ C(d),\ u(a,d)\}$, $\mathcal{T} = \{P_{>0}A \sqsubseteq P_{=1}P_{>0}B,\ C \sqsubseteq P_{>0}(D \sqcap E),\ E \sqsubseteq \exists u.F,\}$ and $\mathcal{S} = \{\exists u.C\}$. Then, applying the rules in Figures 1 and 2 we compute the completed semi-model $\mathbb{M} = \langle \Omega, \mathbb{W}, \tau \rangle$ in the following:*

- $\Omega = \{0,\ P_{>0}A,\ P_{>0}B,\ P_{>0}(D \sqcap E)\}$ and
- $\tau(0) = 0$ and $\tau(\omega) = 1/3$ for all $\omega \in \Omega \setminus \{0\}$.

- $\mathbb{W}(0) = \mathcal{A}^* = \{P_{>0}A(a),\ P_{=1}P_{>0}B(a),\ C(d),\ P_{>0}(D \sqcap E)(d),\ u(a,d),$ $\exists u.C(a),\ P_{>0}B(a)\}$,
- $\mathbb{W}(P_{>0}A) = \{A(a),\ P_{>0}A(a),\ P_{=1}P_{>0}B(a),\ P_{>0}B(a),\ P_{>0}(D \sqcap E)(d)\}$,
- $\mathbb{W}(P_{>0}B) = \{P_{>0}A(a),\ P_{=1}P_{>0}B(a),\ P_{>0}B(a),\ B(a),\ P_{>0}(D \sqcap E)(d)\}$ and
- $\mathbb{W}(P_{>0}(D \sqcap E)) = \{P_{>0}A(a),\ P_{=1}P_{>0}B(a),\ P_{>0}B(a),\ D \sqcap E(d),\ D(d),\ E(d),$

$$\begin{array}{|ll|}\hline P^{+}_{>0}-\text{rule}: & \text{if } C(a)\in\mathbb{W}(\omega) \text{ with } \omega\neq 0,\ P_{>0}C\in\mathcal{C}_{\Sigma,\mathcal{S}},\ \text{and} \\ & P_{>0}C(a)\notin\mathbb{W}(\nu),\ \text{then } \mathbb{W}(\nu):=\mathbb{W}(\nu)\cup\{P_{>0}C(a)\}; \\ P^{-}_{>0}-\text{rule}: & \text{if } P_{>0}C(a)\in\mathbb{W}(\omega) \text{ and } C(a)\notin\mathbb{W}(P_{>0}C), \\ & \text{then } \mathbb{W}(P_{>0}C):=\mathbb{W}(P_{>0}C)\cup\{C(a)\}; \\ P^{+}_{=1}-\text{rule}: & \text{if } \forall\omega\neq 0,\ C(a)\in\mathbb{W}(\omega),\ P_{=1}C\in\mathcal{C}_{\Sigma,\mathcal{S}},\ \text{and} \\ & P_{=1}C(a)\notin\mathbb{W}(\nu),\ \text{then } \mathbb{W}(\nu):=\mathbb{W}(\nu)\cup\{P_{=1}C(a)\}; \\ P^{-}_{=1}-\text{rule}: & \text{if } P_{=1}C(a)\in\mathbb{W}(\omega) \text{ and } C(a)\notin\mathbb{W}(\nu) \text{ with } \nu\neq 0, \\ & \text{then } \mathbb{W}(\nu):=\mathbb{W}(\nu)\cup\{C(a)\}. \\ \hline\end{array}$$

Figure 2: Global expansion rules.

$\exists u.F(d),\ u(d,w^{u}_{F}),\ F(w^{u}_{F}),\ P_{>0}(D\sqcap E)(d)\}.$ □

We denote by Λ the *algorithm* which, given Σ and $\mathcal{C}_{\Sigma,\mathcal{S}}$, nondeterministically applies the expansion rules in Figures 1 and 2 until no further applications are possible. It is easy to see that the size of each set $\mathbb{W}(\omega)$ for $\omega\in\Omega$ is polynomial in $|\Sigma|+|\mathcal{C}_{\Sigma,\mathcal{S}}|$. Since the size of Ω is linear in $|\mathcal{C}_{\Sigma,\mathbb{S}}|$, Λ computes the $\mathbb{M}$ in a time polynomial in $|\Sigma|+|\mathcal{C}_{\Sigma,\mathbb{S}}|$.

Before proving the correctness of Λ, we define the notion of interpretation of a semi-model over Σ, see [12].

Definition 2. *Let $\mathbb{M}=\langle\Omega,\mathbb{W},\tau\rangle$ be a semi-model over Σ, $\mathcal{I}=(\Delta,W,\{\mathcal{I}_x\}_{x\in W},\mu)$ a probabilistic interpretation, and π a mapping from Ω to W. We say that $\mathcal{I}$ satisfies $\mathbb{M}$ via π if for each $\omega\in\Omega$,*

- *$(\mathcal{I},\pi(\omega))\models\mathbb{W}(\omega)$, i.e., $(\mathcal{I},\pi(w))\models\alpha$ for every $\alpha\in\mathbb{W}(\omega)$, and*
- *$\sum_{\omega\in\Omega}\tau(\omega)=\sum_{w\in W}\mu(w)=1$.*

We say that $\mathcal{I}$ satisfies $\mathbb{M}$, denoted as $\mathcal{I}\Vdash\mathbb{M}$, if there is a mapping π such that $\mathcal{I}$ satisfies $\mathbb{M}$ via π.

In the next lemma, we formulate the local soundness property of Λ.

Lemma 1. *Let $\Sigma=\langle\mathcal{A},\mathcal{T}\rangle$ be Prob-$\mathcal{EL}^{>0,=1}_{-\top}$ KB and let $\mathcal{I}=(\Delta,W,\{\mathcal{I}_x\}_{x\in W},\mu)$ be a model of Σ. Also let $\mathbb{M}$ be a semi-model over Σ, α a local or global expansion rule and $\mathbb{M}_\alpha$ a semi-model obtained by applying α to $\mathbb{M}$. If $\mathcal{I}$ satisfies $\mathbb{M}$ via π, then there exists a probabilistic interpretation $\mathcal{I}'=(\Delta,W,\{\mathcal{I}'_x\}_{x\in W},\mu)$ such that*

- *$\mathcal{I}'$ satisfies $\mathbb{M}_\alpha$ via π, and*
- *$\mathcal{I}'$ satisfies Σ.*

Proof. (Outline). We present two cases to illustrate how $\mathcal{I}$ is transformed into $\mathcal{I}'$ by the applications of local and global extension rules; for more details see [16]. Assume the hypotheses and let α be the $\exists^{-}$-rule. Then, for some $\omega\in\Omega$, $\exists r.C(a)\in\mathbb{W}(\omega)$, and since $\mathcal{I}$ satisfies $\mathbb{M}$ via π, we have $(\mathcal{I},\pi(\omega))\models\exists r.C(a)$. By the semantics of existential restriction, there exists a $d\in\Delta$ such that $(a^{\mathcal{I},\pi(\omega)},d)\in r^{\mathcal{I},\pi(\omega)}$ and $d\in C^{\mathcal{I},\pi(\omega)}$. After applying the $\exists^{-}$-rule, $\mathbb{W}(\omega):=\mathbb{W}(\omega)\cup\{r(a,w^{r}_{C}),C(w^{r}_{C})\}$. We have two cases: (1) If w^{r}_{C} occurs in $\mathbb{W}(\omega)$ before the application of the $\exists^{-}$-rule to $\exists r.C(a)$, then $\mathcal{I}'=\mathcal{I}$; (2) If w^{r}_{C} does not occur in $\mathbb{W}(\omega)$ before the application of the $\exists^{-}$-rule to $\exists r.C(a)$, then define the interpretation $(\mathcal{I}',\pi(\omega))$ as $(\mathcal{I},\pi(\omega))$ except for w^{r}_{C}: $(w^{r}_{C})^{\mathcal{I}',\pi(\omega)}=d$. The resulting semi-model $\mathbb{M}_\alpha$ is satisfied by $\mathcal{I}'$ via π. Since $\mathcal{I}$ satisfies Σ and Ω remains unchanged, we conclude that $\mathcal{I}'$ satisfies Σ.

Now let α be the $P^{-}_{>0}$-rule. Then, for some $\omega\in\Omega$, $P_{>0}C(a)\in\mathbb{W}(\omega)$ and $C(a)\notin\mathbb{W}(P_{>0}C)$. Since $\mathcal{I}$ satisfies $\mathbb{M}$ via π, we have $(\mathcal{I},\pi(\omega))\models P_{>0}C(a)$. By the semantics of *Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$, there exists a state $x\in W$ such that $\mu(x)>0$ and $a^{\mathcal{I},x}\in C^{\mathcal{I},x}$. After applying the $P^{-}_{>0}$-rule, $\mathbb{W}(P_{>0}C):=\mathbb{W}(P_{>0}C)\cup\{C(a)\}$.

Now we set $\pi(P_{>0}C) = x$. Then, $\mathbb{M}_\alpha$ is the new semi-model. Hence, $\mathbb{M}_\alpha$ is satisfied by $\mathcal{I}'$ via π where $\mathcal{I}' = \mathcal{I}$. Clearly, $\mathcal{I}'$ satisfies Σ because $\mathcal{I}$ satisfies Σ. □

Lemma 1 makes sure that each application of local and global rules preserves the model existence property. Next we define the *canonical probabilistic interpretation* of a semi-model.

Definition 3. *Let $\mathbb{M}$ be a completed semi-model over Σ. The canonical probabilistic interpretation $\mathcal{I}^c = (\Delta, W, \{\mathcal{I}^c_x\}_{x\in W}, \mu)$ for $\mathbb{M}$ is defined as follows:*

- *$W = \Omega$,*
- *$\mu(0) = \tau(0) = 0$; $\mu(\omega) = \tau(\omega) = 1/ \mid \Omega \setminus \{0\} \mid$ for each $\omega \in \Omega \setminus \{0\}$,*
- *$\Delta = \mathcal{O}^* = \mathcal{O}_\Sigma \cup \mathcal{W}$,*
- *$a^{\mathcal{I}^c,\omega} = a$ for all $a \in \mathcal{O}^*$ and each $\omega \in \Omega$,*
- *$A^{\mathcal{I}^c,\omega} = \{a \in \mathcal{O}^* \mid A(a) \in \mathbb{W}(\omega)\}$, for all $A \in N_C \cap N_\Sigma$,*
- *$r^{\mathcal{I}^c,\omega} = \{(a,b) \in \mathcal{O}^* \times \mathcal{O}^* \mid r(a,b) \in \mathbb{W}(\omega)\}$, for all $r \in N_R \cap N_\Sigma$,*

$(\mathcal{I}^c, \omega)$ is extended to compound concepts in the usual way (see Section 2).

The following lemma shows that $\mathcal{I}^c$ satisfies the completed semi-model $\mathbb{M}$. The proof is by standard induction on the structure of concepts C and hence it is omitted.

Lemma 2. *Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be Prob-$\mathcal{EL}^{>0,=1}_{-\top}$ KB. Also let $\mathbb{M}$ be a completed semi-model over Σ. Then $\mathcal{I}^c \Vdash \mathbb{M}$.*

Next we prove that $(\mathcal{I}^c, \omega) \models \mathcal{T}$, for each $\omega \in \Omega$. We need the following auxiliary lemma whose proof is standard and it is omitted.

Lemma 3. *For each $C \in \mathcal{C}_{\Sigma,\mathcal{S}}$, each $a \in \mathcal{O}^*$ and each $\omega \in \Omega$, if $(\mathcal{I}^c, \omega) \models C(a)$ then $C(a) \in \mathbb{W}(\omega)$.*

Lemma 4. *For each $\omega \in \Omega$, $(\mathcal{I}^c, \omega) \models \mathcal{T}$.*

Proof. Let $\omega \in \Omega$. Suppose that $C \sqsubseteq D \in \mathcal{T}$ and let $a \in C^{\mathcal{I}^c,\omega}$. This means that $(\mathcal{I}^c, \omega) \models C(a)$ and by Lemma 3, $C(a) \in \mathbb{W}(\omega)$. Since $\mathbb{M}$ is completed, by the $\sqsubseteq$-rule, $D(a) \in \mathbb{W}(\omega)$. Since $\mathcal{I}^c \Vdash \mathbb{W}$, by Lemma 2, $(\mathcal{I}^c, \omega) \models D(a)$. Therefore, $(\mathcal{I}^c, \omega) \models C \sqsubseteq D$. Hence, $(\mathcal{I}^c, \omega) \models \mathcal{T}$. □

The following corollary is an immediate consequence of Lemmas 2 and 4.

Corollary 1. *$\mathcal{I}^c$ satisfies Σ.*

Proof. By Definitions 2 and 3 and Lemmas 2 and 4, we have that (1) $(\mathcal{I}^c, 0) \models \Sigma$ and (2) for each $\omega \in \Omega$, $(\mathcal{I}^c, \omega) \models \mathcal{T}$. Hence $\mathcal{I}^c$ satisfies Σ. □

The proof of the next theorem follows immediately from Definition 3 and Lemma 3. In a sense, this theorem captures the completeness property of the algorithm Λ.

Theorem 1. *Let $\mathbb{M}$ be a completed semi-model over Σ and $\mathcal{I}^c = (\Delta, W, \{\mathcal{I}^c_x\}_{x\in W}, \mu)$ a canonical model for $\mathbb{W}$. Then, for all $\omega \in \Omega$, $C \in \mathcal{C}_{\Sigma,\mathcal{S}}$, $r \in N_\Sigma \cap N_R$, and all $a, b \in \mathcal{O}^*$*

- *$(\mathcal{I}^c, \omega) \models r(a,b) \Rightarrow\ r(a,b) \in \mathbb{W}(\omega)$ and*
- *$(\mathcal{I}^c, \omega) \models C(a) \Rightarrow\ C(a) \in \mathbb{W}(\omega)$.*

Finally, the following is a consequence of Theorem 1 and Corollary 1.

Corollary 2. $\mathbb{W}(0) = \mathcal{A}^*$.

4 Secrecy-Preserving Reasoning in *Prob*-$\mathcal{EL}^{>0,=1}_{-\top}$ KBs

Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be a *Prob*-$\mathcal{EL}^{>0,=1}_{-\top}$ KB and $\mathbb{S} \subseteq \mathcal{A}^*$ the "secrecy set". Also let $\mathbb{M}$ be a completed semi-model over Σ. Given Σ, $\mathbb{S}$ and $\mathbb{M}$, the objective is to answer assertion queries while preserving secrecy, i.e., answering queries so that assertions in $\mathbb{S}$ remain protected. Our approach is to compute a function E that assigns a finite set of assertions to each world in Ω. E is called the *secrecy Envelope* for $\mathbb{S}$, so that protecting $E(\omega)$ for all $\omega \in \Omega$, the querying agent cannot logically infer any assertion in $\mathbb{S}$. The sets $E(\omega)$ for each $\omega \in \Omega$ are obtained by applying the *inverted expansion rules* given in Figures 3 and 4. The role of OWA in answering the queries is the following: When answering a query with "Unknown", the querying agent should not be able to distinguish between the case that the answer to the query is truly unknown to the KB reasoner and the case that the answer is being protected for reasons of secrecy. We envision a situation in which once the $\mathbb{W}$ is computed, a *reasoner* $\mathfrak{R}$ is associated with it, i.e., $\mathfrak{R}$ has unfettered access to $\mathbb{W}$. $\mathfrak{R}$ is designed to answer queries as follows: If a query cannot be inferred from Σ, the answer is "Unknown". If it can be inferred and it is not in $E(0)$, the answer is "Yes"; otherwise, the answer is "Unknown". We make the following assumptions about the capabilities of the querying agent:

(a) does not have direct access to ABox $\mathcal{A}$, but is aware of the underlying vocabulary of Σ,
(b) has full access to TBox $\mathcal{T}$,
(c) can ask queries in the form of assertions, and
(d) is not aware of the witness set $\mathcal{W}$, by *hidden name assumptions* (HNA), for more details see [15].

We formally define the notion of an envelope in the following:

Definition 4. *Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be a Prob-$\mathcal{EL}^{>0,=1}_{-\top}$ KB, $\mathbb{S}$ a finite secrecy set and $\mathbb{M}$ a completed semi-model. The secrecy envelope of $\mathbb{S}$ is a function E with domain Ω satisfying the following properties:*

- $\mathbb{S} \subseteq E(0)$,
- *for each $\omega \in \Omega$, $E(\omega) \subseteq \mathbb{W}(\omega)$, and*
- *for each $\omega \in \Omega$, each $\alpha \in E(\omega)$, $\mathbb{W}(\omega) \setminus E(\omega) \not\models \alpha$.*

$$
\begin{array}{rl}
\text{Inv-}\sqcap^{-}\text{–rule}: & \text{if } \{C(a),\ D(a)\} \cap E(\omega) \neq \emptyset \text{ and } C \sqcap D(a)\ \in \mathbb{W}(\omega) \setminus E(\omega), \\
 & \text{then } E(\omega) := E(\omega) \cup \{C \sqcap D(a)\}; \\
\text{Inv-}\sqcap^{+}\text{–rule}: & \text{if } C \sqcap D(a) \in E(\omega),\ \{C(a),\ D(a)\} \subseteq \mathbb{W}(\omega) \setminus E(\omega) \text{ and} \\
 & C \sqcap D \in \mathcal{C}_{\Sigma,\mathcal{S}},\ \text{then } E(\omega) := E(\omega) \cup \{C(a)\} \\
 & \text{or } E(\omega) := E(\omega) \cup \{D(a)\}; \\
\text{Inv-}\exists^{+}\text{–rule}: & \text{if } \exists r.C(a) \in E(\omega),\ \{r(a,b), C(b)\} \subseteq \mathbb{W}(\omega) \setminus E(\omega) \text{ with } b \in \mathcal{O}^* \text{ and} \\
 & \exists r.C \in \mathcal{C}_{\Sigma,\mathcal{S}},\ \text{then } E(\omega) := E(\omega) \cup \{r(a,b)\} \\
 & \text{or } E(\omega) := E(\omega) \cup \{C(b)\}; \\
\text{Inv-}\sqsubseteq\text{–rule}: & \text{if } D(a) \in E(\omega),\ C \sqsubseteq D \in \mathcal{T},\ \text{and } C(a)\ \in \mathbb{W}(\omega) \setminus E(\omega), \\
 & \text{then } E(\omega) := E(\omega) \cup \{C(a)\}.
\end{array}
$$

Figure 3: Inverted local expansion rules.

The intuition for the above definition is that no information in $E(\omega)$ can be inferred from the set $\mathbb{W}(\omega) \setminus E(\omega)$ for each $\omega \in \Omega$. To compute an envelope, we use the idea of inverting the rules of Figures 1 and 2 (see [15], where this approach was first utilized for membership assertions). Induced by the Local and Global expansion rules in Figures 1 and 2, we define the corresponding "inverted" Local and Global expansion rules in Figures 3 and 4, respectively. Note that the $\exists^-$-rule does not have its corresponding inverted rule. The reason for the omission is that an application of this rule results in adding assertions with individual names from the witness set. By HNA, the querying agent is barred from asking any queries that involve individual names from the witness set. Inverted expansion rules are denoted by prefixing Inv- to the name of the corresponding expansion rules.

$$\boxed{\begin{array}{ll}
\text{Inv-}P^{+}_{>0}-\text{rule}: & \text{if } P_{>0}C(a) \in E(\omega),\ P_{>0}C \in \mathcal{C}_{\Sigma,\mathcal{S}},\ \text{and} \\
 & C(a) \in \mathbb{W}(\nu) \setminus E(\nu) \text{ with } \nu \neq 0, \text{ then } E(\nu) := E(\nu) \cup \{C(a)\}; \\
\text{Inv-}P^{-}_{>0}-\text{rule}: & \text{if } C(a) \in E(P_{>0}C) \text{ and } P_{>0}C(a) \notin \mathbb{W}(\omega) \setminus E(\omega), \\
 & \text{then } E(\omega) := E(\omega) \cup \{P_{>0}C(a)\}; \\
\text{Inv-}P^{+}_{=1}-\text{rule}: & \text{if } P_{=1}C(a) \in E(\omega),\ P_{=1}C \in \mathcal{C}_{\Sigma,\mathcal{S}},\ \text{and } \forall \nu \neq 0,\ C(a) \in \mathbb{W}(\nu) \setminus E(\nu), \\
 & \text{then pick a } \tau \neq 0 \text{ such that } E(\tau) := E(\tau) \cup \{C(a)\}; \\
\text{Inv-}P^{-}_{=1}-\text{rule}: & \text{if } C(a) \in E(\omega) \text{ with } \omega \neq 0, \text{ and } P_{=1}C(a) \in \mathbb{W}(\nu) \setminus E(\nu), \\
 & \text{then } E(\nu) := E(\nu) \cup \{P_{=1}C(a)\}.
\end{array}}$$

Figure 4: Inverted global expansion rule.

From now on, we assume that $\mathbb{M}$ has been computed and is readily available for computing the envelope. The computation begins with the initialization step: The set $E(0)$ is initialized as $\mathbb{S}$, and $E(\omega)$ is initialized as $\emptyset$ for all $\omega \in \Omega \setminus \{0\}$. Next, the sets $E(0)$ and $E(\omega)$ for all $\omega \in \Omega \setminus \{0\}$ are expanded using the inverted expansion rules listed in Figures 3 and 4 until no further applications are possible. The resulting function E is said to be completed. We denote by $\Lambda_{\mathbb{S}}$ the algorithm which computes the sets $E(\omega)$ for all $\omega \in \Omega$. Due to non-determinism in applying the rules Inv-$\sqcap^+$ and Inv-$\exists^+$, different executions of $\Lambda_{\mathbb{S}}$ may result different outputs. Since for each $\omega \in \Omega$, $\mathbb{W}(\omega)$ is finite, the computation of $\Lambda_{\mathbb{S}}$ terminates. Let the sets $E(\omega)$ for $\omega \in \Omega$ be an output of $\Lambda_{\mathbb{S}}$. Since the size of each $\mathbb{W}(\omega)$ is polynomial in $|\Sigma| + |\mathcal{C}_{\Sigma,\mathcal{S}}|$, and each application of inverted expansion rule moves an assertion from $\mathbb{W}(\omega)$ into $E(\omega)$, the size of $E(\omega)$ is at most the size of $\mathbb{W}(\omega)$. Since the size of Ω is linear, $\Lambda_{\mathbb{S}}$ may take polynomial time to compute the sets $E(\omega)$. Define the *secrecy-preserving semi-model* for the secrecy set $\mathbb{S}$ to be $\mathbb{M}_E = \langle \Omega, \mathbb{W}_E, \tau \rangle$, where $\mathbb{W}_E(\omega) = \mathbb{W}(\omega) \setminus E(\omega)$ for all $\omega \in \Omega$.

Example 2. *(Example 1 cont.) Recall that $\mathbb{M}_E = \langle \Omega, \mathbb{W}_E, \tau \rangle$ is a completed semi modal. Let $\mathbb{S} = \{P_{=1}P_{>0}B(a),\ P_{>0}(D \sqcap E)(d)\}$ be the secrecy set. Then, applying the rules in Figures 3 and 4 we compute the envelope for $\mathbb{S}$ and one of the corresponding secrecy-preserving semi-model $\mathbb{M}_E = \langle \Omega, \mathbb{W}_E, \tau \rangle$ is given below:*

- $E(0) = \mathbb{S} \cup \{P_{>0}A(a),\ C(d),\ P_{>0}B(a)\}$,
- $E(P_{>0}A) = \{A(a),\ P_{>0}A(a),\ P_{>0}B(a),\ P_{>0}(D \sqcap E)(d)\}$,
- $E(P_{>0}B) = \{B(a), P_{>0}B(a), P_{>0}A(a),\ P_{>0}(D \sqcap E)(d)\}$ and
- $E(P_{>0}(D \sqcap E)) = \{P_{>0}A(a),\ P_{>0}B(a),\ D \sqcap E(d),\ D(d),\ P_{>0}(D \sqcap E)(d)\}$.

- $\mathbb{W}_E(0) = \mathcal{A}^* \setminus E(0) = \{u(a,d), \exists u.C(a)\}$,
- $\mathbb{W}_E(P_{>0}A) = \{P_{=1}P_{>0}B(a)\}$,
- $\mathbb{W}_E(P_{>0}B) = \{P_{=1}P_{>0}B(a)\}$ and
- $\mathbb{W}_E(P_{>0}(D \sqcap E)) = \{P_{=1}P_{>0}B(a),\ E(d),\ \exists u.F(d),\ u(d, w^u_F),\ F(w^u_F)\}$. □

We use this secrecy-preserving semi-model for proving some properties of the envelopes and for answering queries. Before proving the main result on envelopes, we prove several auxiliary lemmas. First, we show that for each $\omega \in \Omega$, no assertions in $E(\omega)$ is "logically reachable" from the members of the set $\mathbb{W}_E(\omega)$.

Lemma 5. *Let the function E be completed by applying the inverted rules in Figures 3 and 4. Also, let $\mathbb{M}_E$ be a secrecy-preserving semi-model. Then, $\mathbb{M}_E$ is completed.*

Proof. Let $\omega \in \Omega$. We have to show that no rule in Figures 1 or 2 is applicable to $\mathbb{W}_E(\omega) = \mathbb{W}(\omega) \setminus E(\omega)$. The proof is by contradiction according to cases: assuming that a rule in Figures 1 and 2 is applicable and showing that some inverse rule is applicable.

- If $\sqcap^-$-rule is applicable, then there is an assertion $C \sqcap D(a) \in \mathbb{W}_E(\omega)$ such that $C(a) \notin \mathbb{W}_E(\omega)$ or $D(a) \notin \mathbb{W}_E(\omega)$. Since $\mathbb{M}$ is completed, $\{C(a), D(a)\} \subseteq \mathbb{W}(\omega)$. Hence, $\{C(a), D(a)\} \cap E(\omega) \neq \emptyset$. This makes the Inv-$\sqcap^-$-rule applicable.

- If $\sqcap^+$-rule is applicable, then there are assertions $C(a), D(a) \in \mathbb{W}_E(\omega)$ such that $C \sqcap D \in \mathcal{C}_{\Sigma,\mathcal{S}}$ and $C \sqcap D(a) \notin \mathbb{W}_E(\omega)$. Since $\mathbb{M}$ is completed, $C \sqcap D(a) \in \mathbb{W}(\omega)$. Hence, $C \sqcap D(a) \in E(\omega)$. This makes the Inv-$\sqcap^+$-rule applicable.

- If $\exists^+$-rule is applicable, then there are assertions $r(a,b), C(b) \in \mathbb{W}_E(\omega)$ such that $\exists r.C \in \mathcal{C}_{\Sigma,\mathcal{S}}$ and $\exists r.C(a) \notin \mathbb{W}_E(\omega)$. Since $\mathbb{M}$ is completed, $\exists r.C(a) \in \mathbb{W}(\omega)$. Hence, $\exists r.C(a) \in E(\omega)$. This makes the Inv-$\exists^+$-rule applicable.

- If $\sqsubseteq$-rule is applicable, then there is an assertion $C(a) \in \mathbb{W}_E(\omega)$ and a GCI $C \sqsubseteq D \in \mathcal{T}$ such that $D(a) \notin \mathbb{W}_E(\omega)$. Since $\mathbb{M}$ is completed, $D(a) \in \mathbb{W}(\omega)$. Hence, $D(a) \in E(\omega)$. This makes the Inv-$\sqsubseteq$-rule applicable.

- If $P^+_{>0}$-rule is applicable, then there is an assertion $C(a) \in \mathbb{W}_E(\omega)$ with $\omega \neq 0$ such that $P_{>0}C \in \mathcal{C}_{\Sigma,\mathcal{S}}$ and $P_{>0}C(a) \notin \mathbb{W}_E(\nu)$ where $\nu \in \Omega$. Since $\mathbb{M}$ is completed, $P_{>0}C(a) \in \mathbb{W}(\nu)$. Hence, $P_{>0}C(a) \in E(\nu)$. This makes the Inv-$P^+_{>0}$-rule applicable.

- If $P^-_{>0}$-rule is applicable, then there is an assertion $P_{>0}C(a) \in \mathbb{W}_E(\omega)$ such that $C(a) \notin \mathbb{W}_E(P_{>0}C)$. Since $\mathbb{M}$ is completed, $C(a) \in \mathbb{W}(P_{>0}C)$. Hence, $C(a) \in E(P_{>0}C)$. This makes the Inv-$P^-_{>0}$-rule applicable.

- If $P^+_{=1}$-rule is applicable, then there is an assertion $C(a) \in \mathbb{W}_E(\omega)$ for each $\omega \neq \Omega \setminus \{0\}$ such that $P_{=1}C \in \mathcal{C}_{\Sigma,\mathcal{S}}$ and $P_{=1}C(a) \notin \mathbb{W}_E(\nu)$ where $\nu \in \Omega$. Since $\mathbb{M}$ is completed, $P_{=1}C(a) \in \mathbb{W}(\nu)$. Hence, $P_{=1}C(a) \in E(\nu)$. This makes the Inv-$P^+_{=1}$-rule applicable.

- If $P^-_{=1}$-rule is applicable, then there is an assertion $P_{=1}C(a) \in \mathbb{W}_E(\omega)$ such that $C(a) \notin \mathbb{W}_E(\nu)$ with $\nu \neq 0$. Since $\mathbb{M}$ is completed, $C(a) \in \mathbb{W}(\nu)$. Hence, $C(a) \in E(\nu)$. This makes the Inv-$P^-_{>0}$-rule applicable.

□

Next we claim that the secrecy-preserving semi-model has similar properties as that of its completed semi-model. The proof is similar to the proofs of the Lemmas 2, 3 and 4.

Lemma 6. *Let $\mathbb{M}_E$ be a secrecy-preserving semi-model obtained from the completed semi-model $\mathbb{M}$ over Σ and the completed function E. Define the canonical probabilistic interpretation $\mathcal{I}^E = (\Delta, W, \{\mathcal{I}^E_x\}_{x \in W}, \mu)$ for $\mathbb{M}_E$ as*

- $W = \Omega$,
- $\mu(0) = \tau(0) = 0$; $\mu(\omega) = \tau(\omega) = 1/\mid \Omega \setminus \{0\} \mid$ *for each* $\omega \in \Omega \setminus \{0\}$,
- $\Delta = \mathcal{O}^* = \mathcal{O}_\Sigma \cup \mathcal{W}$,
- $a^{\mathcal{I}^E,\omega} = a$ *for all* $a \in \mathcal{O}^*$ *and each* $\omega \in \Omega$,
- $A^{\mathcal{I}^E,\omega} = \{a \in \mathcal{O}^* \mid A(a) \in \mathbb{W}_E(\omega)\}$, *for all* $A \in N_C \cap N_\Sigma$,

- $r^{\mathcal{I}^E,\omega} = \{(a,b) \in \mathcal{O}^* \times \mathcal{O}^* \mid r(a,b) \in \mathbb{W}_E(\omega)\}$, *for all* $r \in N_R \cap N_\Sigma$,

$(\mathcal{I}^E, \omega)$ *is extended to compound concepts in the usual way (see Section 2). Then,*

- $\mathcal{I}^E \models\!\!\models \mathbb{M}_E$,
- *For each* $C \in \mathcal{C}_{\Sigma,\mathcal{S}}$*, each* $a \in \mathcal{O}^*$ *and each* $\omega \in \Omega$*, if* $(\mathcal{I}^E, \omega) \models C(a)$*, then* $C(a) \in \mathbb{W}_E(\omega)$ *and*
- *For each* $\omega \in \Omega$*,* $(\mathcal{I}^E, \omega) \models \mathcal{T}$.

Finally, we show that a completed function E is in fact an envelope for the secrecy set $\mathbb{S}$.

Theorem 2. *Let* $\mathbb{M}$ *be a completed semi-model over* Σ*. Also, let* $\mathbb{M}_E$ *be a secrecy-preserving semi-model for the secrecy set* $\mathbb{S}$*. Then, the completed function* E *is an envelope for* $\mathbb{S}$.

Proof. We have to show that the completed function E satisfies all three properties of Definition 4. Properties 1 and 2 are obvious. To prove property 3, suppose that for some $\omega \in \Omega$, some $\alpha \in E(\omega)$, $\mathbb{W}_E(\omega) \models \alpha$.

Let $\mathcal{I}^E = (\Delta, W, \{\mathcal{I}^E_x\}_{x \in W}, \mu)$ be the canonical interpretation for $\mathbb{M}_E$. By Lemma 6, for each $\omega \in \Omega$, $(\mathcal{I}^E, \omega) \models \mathbb{W}_E(\omega)$. Again, by Lemma 6, $\alpha \in \mathbb{W}_E(\omega)$. This is a contradiction. □

5 Query Answering

Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be a *Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$ KB. We assume that the secrecy-preserving semi-model $\mathbb{M}_E$ has been precomputed. The reasoner $\mathfrak{R}$ answers queries based on the information in $\mathbb{M}_E$ and replies to a query q with "Yes" if $\Sigma \models q$ and $q \notin E(0)$; otherwise, the answer is "Unknown". Because of the syntactic restrictions of the language *Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$, $\mathfrak{R}$ does not answer "No" to any query. Since the completed semi-model $\mathbb{M}$ over Σ does not contain all the consequences of Σ, the completed secrecy-preserving semi-model $\mathbb{M}_E$ obtained from $\mathbb{M}$ does not contain all the information needed to answer queries. To address this problems we provide a procedure Eval(k,q) which works by recursively decomposing the compound queries all the way to the information available in $\mathbb{M}_E$. Initial call of this procedure is at the set $\mathbb{W}_E(0)$ of secrecy-preserving semi-model $\mathbb{M}_E$. In lines 1 and 2 of Figure 5, the reasoner checks the membership of q in $\mathbb{W}_E(\omega)$ and answers "Yes" if $q \in \mathbb{W}_E(\omega)$. From line 3 onwards we consider cases in which query q is broken into subqueries based on the constructors defined in *Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$ and apply the procedure recursively. The following theorem states the correctness claim of the algorithm.

Theorem 3. *Let* $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ *be an Prob-*$\mathcal{EL}^{>0,=1}_{-\top}$ *KB,* $\mathbb{M}_E$ *a completed secrecy-preserving semi-model and* q *a query. Then, for every* $\omega \in \Omega$,

- *Soundness: Eval*(ω, q) *outputs "Yes"* $\Rightarrow$ $\mathbb{W}_E(\omega) \models q$;
- *Completeness: Eval*(ω, q) *outputs "Unknown"* $\Rightarrow$ $\mathbb{W}_E(\omega) \not\models q$.

Proof. We omit the proof of soundness. To prove the completeness part assume that $\mathbb{W}_E(\omega) \models q$. We have to show that Eval(ω, q) = "Yes". Let $\mathcal{I}^E$ be the canonical probabilistic interpretation for $\mathbb{M}_E$ as defined in Section 4. By Lamma 6, $\mathcal{I}^E \models\!\!\models \mathbb{M}_E$ and for all $\omega \in \Omega$, $(\mathcal{I}^E, \omega) \models \mathcal{T}$. Therefore, for each $\omega \in \Omega$, $(\mathcal{I}^E, \omega) \models \mathbb{W}_E(\omega)$ and hence, by the assumption, for every ω, $(\mathcal{I}^E, \omega) \models q$. We prove the claim by induction on the structure of q. The inductive hypothesis is, for each $\omega \in \Omega$ and each assertion α if $(\mathcal{I}^E, \omega) \models \alpha$, then Eval$(\omega, \alpha)$ = "Yes". The base case: Let $q = C(a)$ where $C \in \mathcal{C}_{\Sigma,\mathcal{S}}$. Then, by Lemma 6, $C(a) \in \mathbb{W}_E(\omega)$. By Lines 1 and 2 in Figure 5, the claim follows immediately. Next, let $q = C(a)$ where $C \notin \mathcal{C}_{\Sigma,\mathcal{S}}$.

- $q = C \sqcap D(a)$. To answer this query the algorithm computes Eval$(\omega, C(a))$ and Eval$(\omega, D(a))$. Now, the assumption $(\mathcal{I}^E, \omega) \models C \sqcap D(a)$ implies $(\mathcal{I}^E, \omega) \models C(a)$ and $(\mathcal{I}^E, \omega) \models D(a)$ which, by inductive hypothesis, implies that Eval$(\omega, C(a))$ = Eval$(\omega, D(a))$ = "Yes". Hence, by Lines 4 and 5 in Figure 5, Eval$(\omega, C \sqcap D(a))$ = "Yes".

```
Eval($\omega, q$)
 1: case $q \in \mathbb{W}_E(\omega)$
 2:     return "Yes"
 3: case $q = C \sqcap D(a)$
 4:     if Eval($\omega, C(a)$) ="Yes" and Eval($\omega, D(a)$) ="Yes" then
 5:         return "Yes"
 6:     else
 7:         return "Unknown"
 8: case $q = \exists r.C(a)$
 9:     if for some $d \in \mathcal{O}^*$ [ $r(a,d) \in \mathbb{W}_E(\omega)$ and
            Eval($\omega, C(d)$) ="Yes"] then
10:         return "Yes"
11:     else
12:         return "Unknown"
13: case $q = P_{>0}C(a)$
14:     if for some $\nu \in \Omega \setminus \{0\}$ [ Eval($\nu, C(a)$) = "Yes"] then
15:         return "Yes"
16:     else
17:         return "Unknown"
18: case $q = P_{=1}C(a)$
19:     if for each $\nu \in \Omega \setminus \{0\}$ [Eval($\nu, C(a)$) = "Yes"] then
20:         return "Yes"
21:     else
22:         return "Unknown"
```

Figure 5: Query answering algorithm for assertional queries.

- $q = \exists r.C(a)$. By the assumption, $(\mathcal{I}^E, \omega) \models \exists r.C(a)$. This implies that, for some $d \in \mathcal{O}^*, (\mathcal{I}^E, \omega) \models r(a,d)$ and $(\mathcal{I}^E, \omega) \models C(d)$. By Theorem 1, $r(a,d) \in \mathbb{W}_E(\omega)$ and by the inductive hypothesis Eval($\omega, C(d)$)="Yes". Hence, by the Lines 9 and 10 in Figure 5, Eval($\omega, \exists r.C(a)$)= "Yes".
- $q = P_{>0}C(a)$. Then, $(\mathcal{I}^E, \omega) \models P_{>0}C(a)$. This implies that, for some $\nu \in \Omega \setminus \{0\}$, $(\mathcal{I}^E, \nu) \models C(a)$. By the inductive hypothesis Eval($\nu, C(a)$) = "Yes". Hence, by the Lines 13 and 14 in Figure 5, Eval($\omega, P_{>0}C(a)$)= "Yes".
- $q = P_{=1}C(a)$. Then, $(\mathcal{I}^E, \omega) \models P_{=1}C(a)$. This implies that, for each $\nu \in \Omega \setminus \{0\}$, $(\mathcal{I}^E, \nu) \models C(a)$. By the inductive hypothesis Eval($\nu, C(a)$) = "Yes". Hence, by the Lines 18 and 19 in Figure 5, Eval($\omega, P_{=1}C(a)$)= "Yes".

□

Given an assertional query q, the algorithm given in Figure 5 checks for some assertions related to query q in the sets $\mathbb{W}_E(\omega)$ for each $\omega \in \Omega$. Since the size of each set $\mathbb{W}_E(\omega)$ is bounded by $|\Sigma| + |\mathcal{C}_{\Sigma,\mathcal{S}}|$ and also $|\Omega|$ is bounded by $|\Sigma| + |\mathcal{C}_{\Sigma,\mathcal{S}}|$, this algorithm runs in time polynomial in $|\Sigma| + |\mathcal{C}_{\Sigma,\mathcal{S}}|$. Hence the assertional query answering can be done in polynomial time in the size of $|\Sigma| + |\mathcal{C}_{\Sigma,\mathcal{S}}|$.

Example 3. *(example 2 cont.) Recall that $\mathbb{M}_E$ is a secrecy-preserving semi-model. Suppose that the querying agent asks the assertional queries $u(a,d)$, $P_{>0}\exists u.F(a)$, $P_{>0}P_{>0}A(a)$ and $P_{>0}B(a)$. Using the algorithm in Figure 5, we get the following answers:*

q	Eval(ω, q)	Remarks
$u(a,d)$	Yes	by Line 1
$P_{>0}(\exists u.F)(d)$	Yes	by Lines 13 and 8
$P_{>0}P_{>0}A(a)$	Unknown	by Lines 13 and 14
$P_{>0}B(a)$	Unknown	by Line 22 □

6 Conclusions

In this paper we have studied the problem of secrecy-preserving query answering over *Prob*-$\mathcal{EL}^{>0,=1}_{-\top}$ KBs. We have used the conceptual logic-based framework for secrecy-preserving reasoning which was introduced by Tao et al., see [17], to a description logic $\mathcal{EL}$ augmented with the probabilistic constructors $P_{>0}$ and $P_{=1}$. The main contribution is in the way that we compute the consequences and preserve secrecy while answering queries. We break the process into two parts, the first one precomputes the semi-model $\mathbb{M}$ and the envelope E for the given secrecy set $\mathbb{S}$. For this we use two separate (but related) tableau procedures. In query answering step, given $\mathbb{M}$ and E, we define the secrecy-preserving semi-model $\mathbb{M}_E$. Once $\mathbb{M}_E$ has been computed, the query answering procedure is efficient and can be implemented in polynomial time. As for future work, we would like to study secrecy-preserving reasoning framework in temporal description logic $\mathrm{CTL}^{\mathbf{E}\Diamond,\mathbf{E}\Box}_{\mathcal{EL}}$, see [8].

7 Acknowledgments

The first author would like to thank Thurgood Marshall College Fund (TMCF) for their support. The part of the work was done by the first author while visiting at U.S. Army Research Labs (ARL), Adelphi, MD, as a visiting fellow under TMCF summer faculty fellowship.

References

[1] Joachim Biskup and Piero Bonatti. Controlled query evaluation for enforcing confidentiality in complete information systems. *International Journal of Information Security*, 3(1):14–27, 2004.

[2] Joachim Biskup and Piero A Bonatti. Lying versus refusal for known potential secrets. *Data & Knowledge Engineering*, 38(2):199–222, 2001.

[3] Joachim Biskup and Cornelia Tadros. Revising belief without revealing secrets. In *Foundations of Information and Knowledge Systems*, pages 51–70. Springer, 2012.

[4] Joachim Biskup and Torben Weibert. Keeping secrets in incomplete databases. *International Journal of Information Security*, 7(3):199–217, 2008.

[5] Alexandre Evfimievski, Ronald Fagin, and David Woodruff. Epistemic privacy. *Journal of the ACM (JACM)*, 58(1):2, 2010.

[6] Bernardo Cuenca Grau, Evgeny Kharlamov, Egor V Kostylev, and Dmitriy Zheleznyakov. Controlled query evaluation over OWL 2 RL ontologies. In *The Semantic Web–ISWC 2013*, pages 49–65. Springer, 2013.

[7] Bernardo Cuenca Grau, Evgeny Kharlamov, Egor V Kostylev, and Dmitriy Zheleznyakov. Controlled query evaluation over lightweight ontologies. In *Description Logics*, pages 141–152, 2014.

[8] Victor Gutierrez-Basulto, Jean Christoph Jung, and Carsten Lutz. Complexity of branching temporal description logics. In *ECAI*, pages 390–395, 2012.

[9] Joseph Y Halpern. *Reasoning about uncertainty*. MIT press, 2003.

[10] Joseph Y Halpern and Kevin R O'Neill. Secrecy in multiagent systems. *ACM Transactions on Information and System Security (TISSEC)*, 12(1):5, 2008.

[11] Gopalakrishnan Krishnasamy Sivaprakasam and Giora Slutzki. Secrecy-preserving query answering in $\mathcal{ELH}$ knowledge bases. In *Proceedings of 8th International Conference on Agents and Artificial Intelligence*, 2016.

[12] Carsten Lutz and Lutz Schroder. Probabilistic description logics for subjective uncertainty. In *Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning*, pages 393–403. AAAI Press, 2010.

[13] Gopalakrishnan Krishnaswamy Sivaprakasam and Giora Slutzki. Keeping secrets in *EL*ˆ+ knowledge bases. In H. Jaap van den Herik and Joaquim Filipe, editors, *Agents and Artificial Intelligence - 8th International Conference, ICAART 2016, Rome, Italy, February 24-26, 2016, Revised Selected Papers*, volume 10162 of *Lecture Notes in Computer Science*, pages 229–246, 2016.

[14] Gopalakrishnan Krishnaswamy Sivaprakasam and Giora Slutzki. Keeping secrets in modalized DL knowledge bases. In H. Jaap van den Herik, Ana Paula Rocha, and Joaquim Filipe, editors, *Proceedings of the 9th International Conference on Agents and Artificial Intelligence, ICAART 2017, Volume 2, Porto, Portugal, February 24-26, 2017.*, pages 591–598. SciTePress, 2017.

[15] Jia Tao, Giora Slutzki, and Vasant Honavar. Secrecy-preserving query answering for instance checking in $\mathcal{EL}$. In *Proceedings of Web Reasoning and Rule Systems, 195–203*, 2010.

[16] Jia Tao, Giora Slutzki, and Vasant Honavar. Pspace tableau algorithms for acyclic modalized $\mathcal{ALC}$. *Journal of Automated Reasoning*, 49(4):551–582, 2012.

[17] Jia Tao, Giora Slutzki, and Vasant Honavar. A conceptual framework for secrecy-preserving reasoning in knowledge bases. *TOCL*, 16(1):3:1–3:32, 2014.