Inverting Proof Systems for Secrecy under OWA

Giora Slutzki

Department of Computer Science
Iowa State University
Ames, Iowa 50010

slutzki@cs.iastate.edu

May 9th, 2010

Jointly with Jia Tao and Vasant Honavar

Knowledge Representation

- Knowledge representation (KR) mechanisms aim to provide a high level description of a given application domain with the goal of facilitating construction of intelligent applications.
- Representation formalisms based on logic turn out to be eminently suitable because
  1. well-defined syntax
  2. formal semantics
  3. support development of adequate reasoning services

Description Logics

- Description logics (DLs) are a family of logic based Knowledge Representation formalisms.
- DLs describe domain in terms of concepts (classes), roles (binary relationships) and individuals (objects).
  - Decidable fragments of FOL.
  - Closely related to Propositional Modal Logics.
- Formal semantics for DLs are typically model theoretic.

$\mathcal{EL}$ — Concept Expressions and Roles

- Vocabulary: $N_{\mathcal{O}}, N_{\mathcal{C}}, N_{\mathcal{R}}$
- Syntax and semantics: interpretation $\mathcal{I} = (\Delta, \cdot^{\mathcal{I}})$

Syntax	Semantics
$\top$	$\top^{\mathcal{I}} = \Delta$
a	$a^{\mathcal{I}} \in \Delta$
A	$A^{\mathcal{I}} \subseteq \Delta$
r	$r^{\mathcal{I}} \subseteq \Delta \times \Delta$
$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
$\exists r.C$	$\{x \in \Delta \mid \exists y : (x, y) \in r^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$

- Example: $C \sqcap D$, $\exists r.(C \sqcap \exists s.D)$

$\mathcal{EL}$ — Formulae and Knowledge Bases

- $\mathcal{EL}$ formulae are of the form

Syntax	Semantics
$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
$C(a)$	$a^{\mathcal{I}} \in C^{\mathcal{I}}$
$r(a, b)$	$(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$

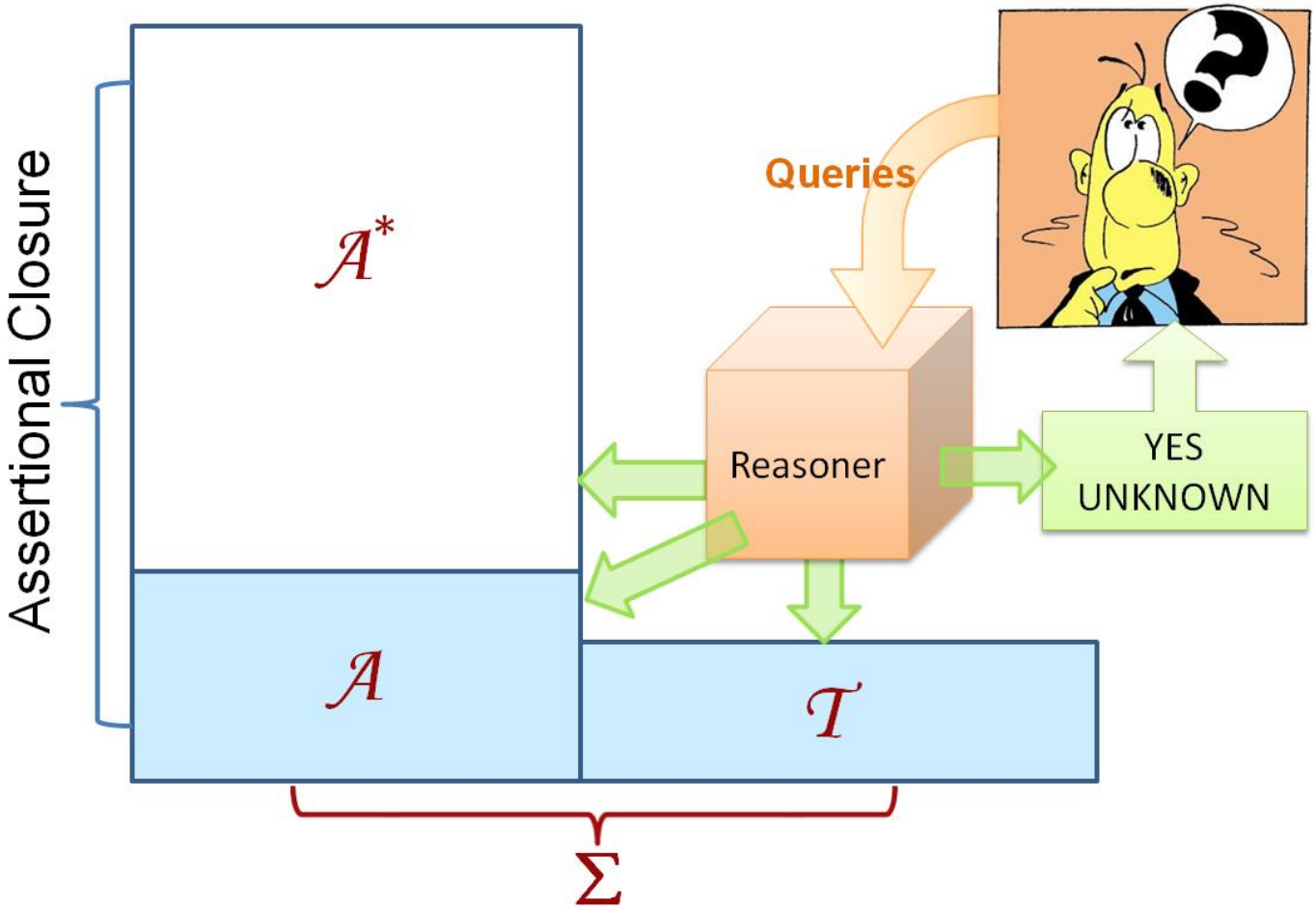
- $\mathcal{EL}$-knowledge base: $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$
  - $\mathcal{A}$: a finite non-empty set of assertions (ABox);
  - $\mathcal{T}$: a finite set of subsumptions (TBox).

DL Reasoning Services

- KB-satisfiability: Σ is satisfiable if it has a model
- Concept-satisfiability: C is satisfiable w.r.t. Σ if there is a model of Σ where the interpretation of C is not empty
- Subsumption: C is subsumed by D w.r.t. Σ if for every model of Σ, the interpretation of C is a subset of that of D
- Query-answering: a is an instance of C if the assertion $C(a)$ is true in every model of Σ

Query Answering

Given a KB $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$, its main goal is to answer user queries. Here we assume that queries are assertions.



Proof System for $\mathcal{A}^*$

$\sqcap_1^{\mathcal{A}}$ -rule:	if $C_1 \sqcap \cdots \sqcap C_k(a) \in \mathcal{A}^*$ and $C_i(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{C_i(a)\}$ where $1 \leq i \leq k$;
$\sqcap_2^{\mathcal{A}}$ -rule:	if $\{C_1(a), ..., C_k(a)\} \subseteq \mathcal{A}^*$, $C_1 \sqcap \cdots \sqcap C_k \in Sub\mathcal{C}$ and $C_1 \sqcap \cdots \sqcap C_k(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{C_1 \sqcap \cdots \sqcap C_k(a)\}$;
$\exists_1^{\mathcal{A}}$ -rule:	if $\{r(a,b), C(b)\} \subseteq \mathcal{A}^*, \exists r.C \in Sub\mathcal{C}$ and $\exists r.C(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{\exists r.C(a)\}$;
$\exists_2^{\mathcal{A}}$ -rule:	if $\exists r.C(a) \in \mathcal{A}^*$ and $\nexists b \in \mathcal{O}^*$ such that $\{r(a,b), C(b)\} \subseteq \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{r(a,c), C(c)\}$ where c is fresh, and $\mathcal{O}^* := \mathcal{O}^* \cup \{c\}$;
$\sqsubseteq^{\mathcal{T}}$ -rule:	if $C(a) \in \mathcal{A}^*, C \sqsubseteq D \in \mathcal{T}$ and $D(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{D(a)\}$.

Theorem: The above proof system is sound and complete.

Query Answering under OWA

Open World Assumption (OWA)

The knowledge of the world is incomplete. Under OWA, if a statement cannot be proven by the reasoner, we do not conclude that it is false. Instead, we view the status of such statements as "Unknown".

Based on OWA, the answer to a query $C(a)$ posed to the knowledge base Σ is defined as

- Yes, if $\Sigma \vdash C(a)$,
- Unknown, otherwise.

Secrecy-preserving Reasoning

OWA: the KB has incomplete information.

Main Idea of Secrecy-preserving Reasoning:

A secrecy-preserving reasoner must answer “Unknown” to every query whose secrecy must be protected. Because of OWA, querying agents are not able to distinguish between the information that is unknown to the reasoner and the information that the reasoner needs to protect.

Goal:

To answer queries as informatively as possible without compromising secret information.

Secrecy Envelopes

Let $\mathbb{S} \subseteq \mathcal{A}^*$ be a set of assertions whose secrecy must be protected.

Secrecy Envelope $\mathbb{E}_{\mathbb{S}}$

$\mathbb{S} \subseteq \mathbb{E}_{\mathbb{S}}$ and $(\mathcal{A}^* \setminus \mathbb{E}_{\mathbb{S}})^* \cap \mathbb{S} = \emptyset$

Tight Envelope $\mathbb{E}^t_{\mathbb{S}}$

$\forall \alpha \in \mathbb{E}^t_{\mathbb{S}}, \quad ((\mathcal{A}^* \setminus \mathbb{E}^t_{\mathbb{S}}) \cup \{\alpha\})^* \cap \mathbb{S} \neq \emptyset.$

Need good algorithms for computing secrecy envelopes.

Example: the knowledge base Σ

$\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$
$\mathcal{T} = \{\exists r.(A \sqcap D) \sqsubseteq C,\ B \sqsubseteq \exists r.D,\ \exists r.D \sqsubseteq C,\ C \sqsubseteq E\}$
$\mathcal{A} = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b)\}$

$\mathcal{T}^*$

- $B \sqsubseteq C,\ B \sqsubseteq E,\ B \sqsubseteq \exists r.D$
- $C \sqsubseteq E$
- $A \sqcap D \sqsubseteq A,\ A \sqcap D \sqsubseteq D$
- $\exists r.(A \sqcap D) \sqsubseteq C,\ \exists r.(A \sqcap D) \sqsubseteq E,\ \exists r.(A \sqcap D) \sqsubseteq \exists r.D$
- $\exists r.D \sqsubseteq C,\ \exists r.D \sqsubseteq E$

$\mathcal{A}^*$

$\mathcal{A} \cup \{A \sqcap D(a), E(a), \exists r.D(a), \exists r.(A \sqcap D)(a)\}$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$

choose $A(a)$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$

$E(a)$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$
$E(a)$, $B(a)$

because $B \sqsubseteq E \in \mathcal{T}^*$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$
$E(a)$, $B(a)$, $C(a)$

because $C \sqsubseteq E \in \mathcal{T}^*$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$
$E(a)$, $B(a)$, $C(a)$, $\exists r.(A \sqcap D)(a)$

because $\exists r.(A \sqcap D) \sqsubseteq E \in \mathcal{T}^*$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$
$E(a)$, $B(a)$, $C(a)$, $\exists r.(A \sqcap D)(a)$, $\exists r.D(a)$

because $\exists r.D \sqsubseteq E \in \mathcal{T}^*$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$
$E(a)$, $B(a)$, $C(a)$, $\exists r.(A \sqcap D)(a)$, $\exists r.D(a)$, $D(a)$

because $\{r(a,a), D(a)\} \subseteq \mathcal{A}^*$ and we choose $D(a)$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$
$E(a)$, $B(a)$, $C(a)$, $\exists r.(A \sqcap D)(a)$, $\exists r.D(a)$, $D(a)$, $r(a,b)$

because $\{r(a,b), D(b)\} \subseteq \mathcal{A}^*$ and we choose $r(a,b)$

Example: a redundant envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_1$

$A \sqcap D(a)$, $A(a)$
$E(a)$, $B(a)$, $C(a)$, $\exists r.(A \sqcap D)(a)$, $\exists r.D(a)$, $D(a)$, $r(a,b)$

$\mathbb{E}_1$ is an envelope. However, $A(a)$ is redundant because of $D(a)$.

Example: a tight envelope

The secrecy set $\mathbb{S} = \{A \sqcap D(a), E(a)\}$

$\mathcal{A}^* = \{A(a), B(a), D(a), C(a), r(a,a), r(a,b), D(b), A \sqcap D(a), E(a),$
$\exists r.D(a), \exists r.(A \sqcap D)(a)\}$

The secrecy envelope $\mathbb{E}_2$

$A \sqcap D(a)$, $D(a)$,
$E(a)$, $B(a)$, $C(a)$, $\exists r.(A \sqcap D)(a)$, $\exists r.D(a)$, $r(a,b)$

$\mathbb{E}_2$ is tight.

Computing Secrecy Envelopes

How to compute secrecy envelopes that are both:

- informative, and
- secrecy-preserving.

Tight would be good! Optimal would be better, but

Computing Secrecy Envelopes

How to compute secrecy envelopes that are both:

- informative, and
- secrecy-preserving.

Tight would be good! Optimal would be better, but

The Secrecy Envelope Problem is NP-complete

Given a KB $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ and a secrecy set $\mathbb{S} \subseteq \mathcal{A}^*$, let $k \leq |\mathcal{A}^*|$. Is there a secrecy envelope $\mathbb{E}$ such that $\mathbb{S} \subseteq \mathbb{E} \subseteq \mathcal{A}^*$ and $|\mathbb{E} \setminus \mathbb{S}| \leq k$?

Computing Secrecy Envelopes

How to compute secrecy envelopes that are both:

- informative, and
- secrecy-preserving.

Lazy approach:

wait for queries; when query α comes along, figure out how to answer it so that no information about secrecy set $\mathbb{S}$ is revealed, taking into account answers to prior queries:

$$(Q_{YES} \cup \{\alpha\})^* \cap \mathbb{S} = \emptyset$$

Main Idea

Take the reasoner's proof system used to compute consequences $\mathcal{A}^*$ of the KB $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ and "invert" it into a "proof system" to compute the secrecy envelope $\mathbb{E}_{\mathbb{S}}$ from the secrecy set $\mathbb{S}$.

Approach

We invert the inference rules.

Illustrations (non $\mathcal{EL}$)

1. Modus Ponens

$$\frac{A, \quad A \to B}{B}$$

2. And-Elimination

$$\frac{A \wedge B}{A, B}$$

3. And-Introduction

$$\frac{A, B}{A \wedge B}$$

1. Inverse Modus Ponens

$$\frac{\text{B is secret}, \quad A \to B}{A \text{ should be secret}}$$

2. Inverse And-Elimination

$$\frac{A \wedge B \text{ is secret}}{A \text{ or } B \text{ should be secret}}$$

3. Inverse And-Introduction

$$\frac{A \text{ or } B \text{ is secret}}{A \wedge B \text{ should be secret}}$$

$\mathcal{EL}$ secrecy closure rules

$\sqcap_1$ rules:

$\sqcap_1^{\mathcal{A}}$-rule: If $C_1 \sqcap \cdots \sqcap C_k(a) \in \mathcal{A}^*$ and $C_i(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{C_i(a)\}$ where $1 \leq i \leq k$

$\sqcap_1^{\mathbb{S}}$-rule: If $C_1 \sqcap \cdots \sqcap C_k(a) \in \mathcal{A}^* \setminus \mathbb{E}$ and $\{C_1(a), ..., C_k(a)\} \cap \mathbb{E} \neq \emptyset$, then $\mathbb{E} := \mathbb{E} \cup \{C_1 \sqcap \cdots \sqcap C_k(a)\}$

$\sqcap_2$ rules:

$\sqcap_2^{\mathcal{A}}$-rule: If $\{C_1(a), ..., C_k(a)\} \subseteq \mathcal{A}^*$, $C_1 \sqcap \cdots \sqcap C_k \in \mathit{Sub}\mathcal{C}$ and $C_1 \sqcap \cdots \sqcap C_k(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{C_1 \sqcap \cdots \sqcap C_k(a)\}$

$\sqcap_2^{\mathbb{S}}$-rule: If $C_1 \sqcap \cdots \sqcap C_k(a) \in \mathbb{E}$ and $\{C_1(a), ..., C_k(a)\} \cap \mathbb{E} = \emptyset$, then $\mathbb{E} := \mathbb{E} \cup \{C_i(a)\}$ where $1 \leq i \leq k$

$\mathcal{EL}$ secrecy closure rules

$\exists_1$ rules:

$\exists_1^{\mathcal{A}}$-rule: If $\{r(a,b), C(b)\} \subseteq \mathcal{A}^*, \exists r.C \in Sub\mathcal{C}$ and $\exists r.C(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{\exists r.C(a)\}$

$\exists_1^{\mathbb{S}}$-rule: If $\exists r.C(a) \in \mathbb{E}, \exists b \in \mathcal{O}^*$ s.t. $\{r(a,b), C(b)\} \subseteq \mathcal{A}^* \setminus \mathbb{E}$, then $\mathbb{E} := \mathbb{E} \cup \{r(a,b)\}$ or $\mathbb{E} := \mathbb{E} \cup \{C(b)\}$

$\exists_2$ rules:

$\exists_2^{\mathcal{A}}$-rule: If $\exists r.C(a) \in \mathcal{A}^*$ and $\nexists b \in \mathcal{O}^*$ such that $\{r(a,b), C(b)\} \subseteq \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{r(a,c), C(c)\}$ where c is fresh, and $\mathcal{O}^* := \mathcal{O}^* \cup \{c\}$

$\exists_2^{\mathbb{S}}$-rule: If $\exists r.C(a) \in \mathcal{A}^* \setminus \mathbb{E}$ and $\forall b \in \mathcal{O}^*$ with $\{r(a,b), C(b)\} \subseteq \mathcal{A}^*$, we have $\{r(a,b), C(b)\} \cap \mathbb{E} \neq \emptyset$, then $\mathbb{E} := \mathbb{E} \cup \{\exists r.C(a)\}$

$\mathcal{EL}$ secrecy closure rules

$\sqsubseteq$ rules:

$\sqsubseteq^{\mathcal{T}}$-rule: If $C(a) \in \mathcal{A}^*, C \sqsubseteq D \in \mathcal{T}$ and $D(a) \notin \mathcal{A}^*$, then $\mathcal{A}^* := \mathcal{A}^* \cup \{D(a)\}$

$\sqsubseteq^{\mathbb{S}}$-rule: If $D(a) \in \mathbb{E}, C \sqsubseteq D \in \mathcal{T}$ and $C(a) \in \mathcal{A}^* \setminus \mathbb{E}$, then $\mathbb{E} := \mathbb{E} \cup \{C(a)\}$

Theorem.

Let $\Sigma = \langle \mathcal{A}, \mathcal{T} \rangle$ be a knowledge base, $\mathbb{S} \subseteq \mathcal{A}^*$ a secrecy set and let $\mathbb{E}$ be obtained from $\mathbb{S}$ by the secrecy closure rules until none is applicable. Then $\mathbb{E}$ is a secrecy envelope of $\mathbb{S}$.

Remark: The envelope $\mathbb{E}$ may not be tight.

Computing Tight Envelopes

1. Deterministic version of $\exists^{\mathbb{S}}_1$-rule:

   $\exists^{\mathbb{S}}_{1d}$-rule: if $\exists r.C(a) \in \mathbb{E}, \exists b \in \mathcal{O}^*$ s.t. $\{r(a,b), C(b)\} \subseteq \mathcal{A}^* \setminus \mathbb{E}$, then $\mathbb{E} := \mathbb{E} \cup \{r(a,b)\}$.

2. Drop $\exists^{\mathbb{S}}_2$-rule:

   $\exists^{\mathbb{S}}_2$-rule: if $\exists r.C(a) \in \mathcal{A}^* \setminus \mathbb{E}$, and $\forall b \in \mathcal{O}^*$ with $\{r(a,b), C(b)\} \subseteq \mathcal{A}^*$, we have $\{r(a,b), C(b)\} \cap \mathbb{E} \neq \emptyset$, then $\mathbb{E} := \mathbb{E} \cup \{\exists r.C(a)\}$

3. Apply remaining secrecy closure rules in a specific order while removing redundancy.

Computing Tight Envelopes

We show that

- The set $\mathbb{E}$, $\mathbb{S} \subseteq \mathbb{E}$, resulting from this process is a tight secrecy envelope of $\mathbb{S}$, and
- $\mathbb{E}$ can be computed in polynomial time.

Secrecy-Preserving Query Answering

SPQA($\mathcal{T}, \mathcal{A}^*, C(a), \mathbb{E}_{\mathbb{S}}$):
1. if ($C \notin Sub\mathcal{C}$)
2. {
3. compute $sub(C)$;
4. update $\mathcal{A}^*$ by adding the concepts in $sub(C) \setminus Sub\mathcal{C}$
5. expand the secrecy envelope $\mathbb{E}_{\mathbb{S}}$
6. }
7. if ($C(a) \in \mathcal{A}^*$ and $C(a) \notin \mathbb{E}_{\mathbb{S}}$)
8. return "Yes"
9. else
10. return "Unknown"

Figure: Secrecy Preserving Query Answering procedure

Thank you!