Unification in the Description Logic $\mathcal{EL}$

Franz Baader and Barbara Morawska

TU Dresden, Germany

UNIF 2009

UNIF 2008 Unification in $\mathcal{EL}$ is of type zero.

UNIF 2009 Unification in $\mathcal{EL}$ is decidable and is in NP.
Unification problem in $\mathcal{EL}$ is NP-complete.

Outline

Description Logic $\mathcal{EL}$

- Concept names: City, Cathedral,
- Top concept: $\top$,
- Conjunction: $\sqcap$,
- Existential restriction: $\exists$*has-location*.$\top$

Example (concept term)

City $\sqcap$ $\exists$ *location. East-South of Germany* $\sqcap$ $\exists$ *university.* $\top$

Description Logic $\mathcal{EL}$

Semantics

$(\Delta, \mathcal{I})$ is an interpretation, where:

- Concepts are sets: if $A \in N_C$, $A^{\mathcal{I}} \subseteq \Delta$;
- Roles are binary relations:if $r \in N_R$, $r^{\mathcal{I}} \subseteq \Delta \times \Delta$;
- $\top$ is the domain: $\top^{\mathcal{I}} = \Delta$;
- Conjunction is intersection: $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$;
- $(\exists r.C)^{\mathcal{I}} = \{c \in \Delta \mid \exists b \in \Delta.(c, b) \in r^{\mathcal{I}} \text{and } b \in C^{\mathcal{I}}\}$

Subsumption and equivalence

- Subsumption:
  $C \sqsubseteq D$ iff for all interpretations $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.
- Equivalence:
  $C \equiv D$ iff $C \sqsubseteq D$ and $D \sqsubseteq C$

Variables in $\mathcal{EL}$

We define a set of variables N_V as a subset of N_C.

Idea: concept names in N_V may be defined differently by different users or developers of a given ontology.

Concepts from N_V can be substituted with concept terms, concepts from N_C cannot be substituted.

$\mathcal{EL}$-Unification

Example:

- $\mathit{City} \sqcap \exists\, \mathit{location}.\ \mathit{East\text{-}South\ of\ Germany} \sqcap \exists\, \mathit{size}.\ (\mathit{more\text{-}than\text{-}500000} \sqcap \mathit{less\text{-}than\text{-}1000000})$
- $\mathit{Settlement} \sqcap \exists\, \mathit{has}.\ \mathit{Cathedral} \sqcap \exists\, \mathit{location}.\mathit{Saxony} \sqcap \exists\, \mathit{size}.\ \mathit{middle}$

$\mathcal{EL}$-Unification

$\mathcal{EL}$-Unification Problem

is a set of equalities, $C_1 \equiv^? D_1, \ldots, C_n \equiv^? D_n$, where C_i, D_i are $\mathcal{EL}$-concept terms.

A substitution σ is an $\mathcal{EL}$-unifier (solution)

of an $\mathcal{EL}$-unification problem $C_1 \equiv^? D_1, \ldots, C_n \equiv^? D_n$ if $\sigma(C_1) \equiv \sigma(D_1), \ldots, \sigma(C_n) \equiv \sigma(D_n)$.

$SLmO$ – semilattices with monotone operators

$$SLmO = \{ x \wedge (y \wedge z) = (x \wedge y) \wedge z,\quad x \wedge y = y \wedge z,\quad x \wedge x = x,\quad x \wedge 1 = x,\quad \{f_i(x \wedge y) \wedge f_i(y) = f_i(x \wedge y) \mid 1 \leqslant i \leqslant n\} \}$$

- $\sqcap$ is associative, commutative and idempotent,
- $\top$ is a unit for $\sqcap$
- $\exists r_i.(C \sqcap D) \sqcap \exists r_i.D \equiv \exists r_i.(C \sqcap D)$

Existential restriction is not a homomorphism:
$\exists r.(A \sqcap B) \subsetneq \exists r.A \sqcap \exists r.B$

$\mathcal{EL}$-problem of Type Zero

What are the unifiers of the following goal:

$\exists R.Y \sqsubseteq^? X$

For example:

- $[X \mapsto \exists R.Z_1, \quad Y \mapsto Z_1]$
- $[X \mapsto \exists R.Z_1 \sqcap \exists R.Z_2, \quad Y \mapsto Z_1 \sqcap Z_2]$
- $[X \mapsto \exists R.Z_1 \sqcap \exists R.Z_2 \sqcap \exists R.Z_3, \quad Y \mapsto Z_1 \sqcap Z_2 \sqcap Z_3]$
- ...

Reductions and reduced forms in $\mathcal{EL}$

Reduction rules are applied to concept terms modulo $\mathcal{AC}$

- $C \sqcap \top \rightsquigarrow C$
- $A \sqcap A \rightsquigarrow A$
- if $D \sqsubseteq C$, then $\exists r.D \sqcap \exists r.C \rightsquigarrow \exists r.D$

Equivalence of reduced concepts

Theorem (Küsters)

$$C \equiv D \qquad \textit{iff} \qquad \widehat{C} =_{\mathcal{AC}} \widehat{D}$$

$$\textit{where } C \rightsquigarrow \widehat{C},\ D \rightsquigarrow \widehat{D}$$

Inverse of subsumption

Subsumption order: $C_1 > C_2$ iff $C_1 \sqsupset C_2$.
Subsumption order is not well founded.

Inverse of subsumption order: $C_1 >_{is} C_2$ iff $C_1 \sqsubset C_2$.

Lemma

There is no infinite sequence $C_0, C_1, C_2, \ldots$ of $\mathcal{EL}$-concept terms such that $C_0 \sqsubset C_1 \sqsubset C_2 \sqsubset \cdots$.

Monotonicity of $>_{is}$

Lemma

C is a reduced concept term and contains D,

$$D >_{is} D'$$

Then:

$$C >_{is} C'$$

where C' is obtained from C by relpalcing an occurrence of D by D'.

Proof

Induction on size of C.

1. $C = D$, obvious.
2. $C = \exists R.C_1$ and D occurs in C_1 (induction).
3. $C = C_1 \sqcap \cdots \sqcap C_n$ and D occurs in C_i.

Monotonicity of $>_{is}$

Proof of the case where $C = C_1 \sqcap \cdots \sqcap C_n$ and D occurs in C_1.

$$C_1 \sqcap \cdots \sqcap C_n \leadsto C_1' \sqcap C_2 \sqcap \cdots \sqcap C_n$$

By induction $C_1 >_{is} C_1'$, i.e. $C_1 \sqsubset C_1'$.
and by monotonicity of $\sqsubseteq$:
$C_1 \sqcap \cdots \sqcap C_n \sqsubseteq C_1' \sqcap C_2 \sqcap \cdots \sqcap C_n$
Hence
$C_1 \sqcap \cdots \sqcap C_n \not>_{is} C_1' \sqcap C_2 \sqcap \cdots \sqcap C_n$
means $C_1 \sqcap \cdots \sqcap C_n \equiv C_1' \sqcap C_2 \sqcap \cdots \sqcap C_n$
$C_1 \not\equiv C_1'$, there is $i \neq 1$, such that

$$C_1 \sqsubset C_1' \equiv C_i.$$

But this means that C_1 "eats up" C_i in C, and thus C is not reduced. Contradiction.

Minimal unifiers

$>_{is}$ is well-founded
its multiset extension $>_m$ is well-founded.

$S(\sigma)$ as a multiset of all $\sigma(X)$, $X \in Var(\Gamma)$.

Definition

$\sigma \succ \gamma$ iff $S(\sigma) >_m S(\gamma)$.
σ, θ are ground, reduced unifiers of Γ.

The ground, reduced unifier σ of Γ is minimal iff there is no unifer θ, such that $\sigma \succ \theta$.

Obviously, a goal is unifiable iff it has a minimal ground reduced unifier.

Atoms and flat goals

A concept term is an atom iff it is a constant or of form $\exists r.C$.

A flat atom is an atom which is a *constant* or $\exists r.C$, where C *is constant, variable or* $\top$.

A goal Γ is flat iff it only contains the equations of the form:

- $X \equiv^? C$ with X a variable and C a non-variable flat atom,
- $X_1 \sqcap \cdots \sqcap X_m \equiv^? Y_1 \sqcap \cdots \sqcap Y_n$,
  where $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ are variables.

Atoms of a unifier σ

$$At(\sigma) = \bigcup_{X \in Var(\Gamma)} At(\sigma(X))$$

Definition

For every concept term C, define $At(C)$:

- if $C = \top$, then $At(C) = \emptyset$,
- if C is a constant, then $At(C) = \{C\}$,
- if $C = \exists r.D$, then $At(C) = \{C\} \cup At(D)$,
- if $C = D_1 \sqcap D_2$, then $At(C) = At(D_1) \cup At(D_2)$.

Locality of a minimal ground reduced unifier

γ is a minimal reduced ground unifier of Γ

Lemma

If C is an atom of γ,
then there is a non-variable atom D in Γ,
such that $C \equiv \gamma(D)$

Proof by contradiction.

Idea: If C is maximal w. r. t. $\sqsubseteq$ and violates the lemma, we construct a smaller unifier γ' – contradiction.

- C is a constant A.
- C is of the form $\exists r.C_1$.

Proof of the case where C is of the form $\exists r.C_1$

$D_1, \ldots, D_n$ are all atoms in Γ, such that
$C \sqsubset \gamma(D_1), \ldots, C \sqsubset \gamma(D_n)$.

$$C \sqsubset \gamma(D_1) \sqcap \cdots \sqcap \gamma(D_n).$$

Obtain γ' by replacing C with reduced form of
$\gamma(D_1) \sqcap \cdots \sqcap \gamma(D_n)$.

Check if γ' is also a unifier of Γ

- $X \equiv^? E$,
- $X_1 \sqcap \cdots \sqcap X_m \equiv^? Y_1 \sqcap \cdots \sqcap Y_n$,

$$\gamma(X_1) \sqcap \cdots \sqcap \gamma(X_m) \equiv \gamma(Y_1) \sqcap \cdots \sqcap \gamma(Y_n)$$
$$\gamma(X_1) \sqcap \cdots \sqcap \gamma(X_m) \rightsquigarrow [U]_{\mathcal{AC}} \leftsquigarrow \gamma(Y_1) \sqcap \cdots \sqcap \gamma(Y_n)$$

We show that all these reductions are preserved
if C is replaced by reduced $\gamma(D_1) \sqcap \cdots \sqcap \gamma(D_n)$.

The most interesting reduction is:

$$\exists r.E_1 \sqcap \exists r.E_2 \rightsquigarrow \exists r.E_1$$

if $E_1 \sqsubseteq E_2$

Assume that C is in E_1 and there is C' in E_2, such that $C \sqsubseteq C'$.

- $C = C'$, (easy, both are replaced by $\gamma(D_1) \widehat{\sqcap \cdots \sqcap} \gamma(D_n)$),
- $C \sqsubset C'$

In the second case $C' = \top$ or C' is $\gamma(D_i)$, and
$\gamma(D_1) \sqcap \cdots \sqcap \gamma(D_n) \sqsubset C'$.

Corollary

Corollary

Γ – a flat goal
γ – minimal reduced ground unifier of Γ
$X \in Var(\Gamma)$

Then $\gamma(X) = \top$ or there are non-variable atoms $D_1, \ldots, D_n$ ($n \geqslant 1$) of Γ such that $\gamma(X) \equiv \gamma(D_1) \sqcap \cdots \sqcap \gamma(D_n)$.

Algorithm

Algorithm

1. For each X in Γ guess a set S_X of non-variable atoms in Γ.
2. Define: X depends on Y if Y occurs in S_X.
   Fail if there are circular dependencies in the transitive closure of *depends*.
3. Define a substitution
   - If S_X is empty, then $\sigma(X) = \top$,
   - otherwise, $S_X = \{D_1, \ldots, D_n\}$ and $\sigma(X) = \sigma(D_1) \sqcap \cdots \sqcap \sigma(D_n)$.
4. Check if σ is a unifier of Γ.

Complexity

Theorem

$\mathcal{EL}$-unification is NP-complete.

Proof.

The problem is *NP*-hard, because $\mathcal{EL}$-matching is *NP*-hard.

Consider the algorithm:
Present the subsitution σ as a sequence of equations, a TBox T_σ. (Hence the definition of σ is polynomial)

For each $C \equiv^? D \in \Gamma$, $\sigma(C) \equiv \sigma(D)$ iff $C \equiv_{T_\sigma} D$.

In $\mathcal{EL}$ subsumption (and thus equivalence) modulo acyclic TBoxes is polynomial. □

(What is a minimal unifier of the "type-zero" problem?)

Conclusion

We have shown

Unification in $\mathcal{EL}$ is *NP*-complete.

What next?

- Implementation...
- Unification in $\mathcal{EL}$ but without $\top$...
- Unification in extensions of $\mathcal{EL}$, e.g. $\forall r.C$.