Quick Introduction to Description Logic

In Description Logic (DL), we distinguish between

- conceptual or terminological knowledge, stored in an ontology $\mathcal{O}$ (often also called TBox $\mathcal{T}$); and
- data, stored in a data instance $\mathcal{D}$ (often also called an ABox $\mathcal{A}$).

.

Quick Introduction to Description Logic

In Description Logic (DL), we distinguish between

- conceptual or terminological knowledge, stored in an ontology $\mathcal{O}$ (often also called TBox $\mathcal{T}$); and
- data, stored in a data instance $\mathcal{D}$ (often also called an ABox $\mathcal{A}$).

A pair $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ is a knowledge base (KB).

Quick Introduction to Description Logic

In Description Logic (DL), we distinguish between

- conceptual or terminological knowledge, stored in an ontology $\mathcal{O}$ (often also called TBox $\mathcal{T}$); and
- data, stored in a data instance $\mathcal{D}$ (often also called an ABox $\mathcal{A}$).

A pair $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ is a knowledge base (KB).

$\mathcal{D}$ is a finite set of assertions of the form $A(b)$, $R(a, b)$ with

- A a concept name (unary relation);
- R a role name (binary relation);
- and a, b individual names.

Quick Introduction to Description Logic

In Description Logic (DL), we distinguish between

- conceptual or terminological knowledge, stored in an ontology $\mathcal{O}$ (often also called TBox $\mathcal{T}$); and
- data, stored in a data instance $\mathcal{D}$ (often also called an ABox $\mathcal{A}$).

A pair $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ is a knowledge base (KB).

$\mathcal{D}$ is a finite set of assertions of the form $A(b)$, $R(a, b)$ with

- A a concept name (unary relation);
- R a role name (binary relation);
- and a, b individual names.

$\mathcal{O}$ is a finite set of concept inclusions (CIs) of the form $C \sqsubseteq D$ with C, D concepts from a DL $\mathcal{L}$.

Description Logic $\mathcal{EL}$

$\mathcal{EL}$-concepts are constructed according to

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

Description Logic $\mathcal{EL}$

$\mathcal{EL}$-concepts are constructed according to

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

Typically used for large ontologies (for instance SNOMED CT).

Description Logic $\mathcal{EL}$

$\mathcal{EL}$-concepts are constructed according to

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

Typically used for large ontologies (for instance SNOMED CT).

Some concept inclusions:

$$\begin{array}{rcl} \exists \text{support.FootballClub} & \sqsubseteq & \text{Footballfan} \\ \text{FootballClub} & \sqsubseteq & \text{Club} \sqcap \exists \text{competes\_in.FootballLeague} \end{array}$$

Description Logic $\mathcal{EL}$

$\mathcal{EL}$-concepts are constructed according to

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

Typically used for large ontologies (for instance SNOMED CT).

Some concept inclusions:

$$\begin{array}{rcl} \exists \text{support.FootballClub} & \sqsubseteq & \text{Footballfan} \\ \text{FootballClub} & \sqsubseteq & \text{Club} \sqcap \exists \text{competes\_in.FootballLeague} \end{array}$$

Data instance:

FootballClub(LiverpoolFC),
competes_in(LiverpoolFC, PremierLeague),
FootballLeague(PremierLeague)

Description Logic $\mathcal{ALC}$

$\mathcal{ALC}$-concepts are constructed according to

$$C, D := A \mid \top \mid C \sqcap D \mid \neg C \mid \exists R.C$$

We set $\forall R.C := \neg\exists R.\neg C$ and $\bot = \neg\top$.

Description Logic $\mathcal{ALC}$

$\mathcal{ALC}$-concepts are constructed according to

$$C, D := A \mid \top \mid C \sqcap D \mid \neg C \mid \exists R.C$$

We set $\forall R.C := \neg \exists R.\neg C$ and $\bot = \neg \top$.

Step towards expresssive DLs underpinning OWL standard.

Description Logic $\mathcal{ALC}$

$\mathcal{ALC}$-concepts are constructed according to

$$C, D := A \mid \top \mid C \sqcap D \mid \neg C \mid \exists R.C$$

We set $\forall R.C := \neg\exists R.\neg C$ and $\bot = \neg\top$.

Step towards expresssive DLs underpinning OWL standard.

Some concept inclusions:

$$\begin{array}{rcl} \text{Footballfan} & \sqsubseteq & \neg\text{Cricketfan} \\ \text{FootballClub} & \sqsubseteq & \forall\text{competes\_in.FootballLeague} \end{array}$$

Description Logic $\mathcal{ALC}$

$\mathcal{ALC}$-concepts are constructed according to

$$C, D := A \mid \top \mid C \sqcap D \mid \neg C \mid \exists R.C$$

We set $\forall R.C := \neg \exists R.\neg C$ and $\bot = \neg \top$.

Step towards expresssive DLs underpinning OWL standard.

Some concept inclusions:

$$\begin{array}{rcl} \text{Footballfan} & \sqsubseteq & \neg\text{Cricketfan} \\ \text{FootballClub} & \sqsubseteq & \forall\text{competes\_in.FootballLeague} \end{array}$$

Sometimes we also use inverse roles, R^{-}.

$$\text{FootballLeague} \sqsubseteq \forall \text{competes\_in}^{-}.\text{FootballClub}$$

The resulting languages are denoted $\mathcal{ELI}$ and $\mathcal{ALCI}$, respectively.

Reasoning

Description logics are interpreted in interpretations

$$\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$$

with $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, and $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, in the expected way. We write $C^{\mathcal{I}}$ for the inductively defined extension of C in $\mathcal{I}$ and use

$$\mathcal{O} \models C \sqsubseteq D, \quad (\mathcal{O}, \mathcal{D}) \models R(a, b), \quad (\mathcal{O}, \mathcal{D}) \models A(a)$$

as in first-order logic. Open world semantics!

Reasoning

Description logics are interpreted in interpretations

$$\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$$

with $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, and $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, in the expected way. We write $C^{\mathcal{I}}$ for the inductively defined extension of C in $\mathcal{I}$ and use

$$\mathcal{O} \models C \sqsubseteq D, \quad (\mathcal{O}, \mathcal{D}) \models R(a, b), \quad (\mathcal{O}, \mathcal{D}) \models A(a)$$

as in first-order logic. Open world semantics!

Deciding these entailments is

- in PTime for $\mathcal{EL}$;
- ExpTime-complete for $\mathcal{ELI}$, $\mathcal{ALC}$, $\mathcal{ALCI}$.

The Separability Problem

Let $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ and $E = (E^+, E^-)$ with

- $E^+ \subseteq \text{ind}(\mathcal{D})$ a set of positive example
- $E^- \subseteq \text{ind}(\mathcal{D})$ a set of negative examples,

A concept C (sometimes also formula or query C) separates E under $\mathcal{K}$ if it applies to all $a \in E^+$ but not to any $a \in E^-$.

The Separability Problem

Let $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ and $E = (E^+, E^-)$ with

- $E^+ \subseteq \text{ind}(\mathcal{D})$ a set of positive example
- $E^- \subseteq \text{ind}(\mathcal{D})$ a set of negative examples,

A concept C (sometimes also formula or query C) separates E under $\mathcal{K}$ if it applies to all $a \in E^+$ but not to any $a \in E^-$.

We aim to determine and investigate a few important dimensions of separation and the problem of deciding separability.

We do not yet look into the problem learning C from E in the sense of finding some good generalisation of E.

Examples

Let $\mathcal{K}_1 = (\emptyset, \mathcal{D})$ where

$$\begin{aligned}\mathcal{D} \quad = \quad & \{\text{citizen\_of}(\text{Peter}, \text{UK}), \text{citizen\_of}(\text{Piotr}, \text{Poland}), \\ & \text{citizen\_of}(\text{Kazue}, \text{Japan}), \\ & \text{EuropC}(\text{UK}), \text{EuropC}(\text{Poland}), \\ & \text{Person}(\text{Peter}), \text{Person}(\text{Piotr})\}.\end{aligned}$$

Let $E^+ = \{\text{Peter}, \text{Piotr}\}$ and $E^- = \{\text{Kazue}\}$.

Task: Separate E using an $\mathcal{ALCI}$-concept.

Examples

Let $\mathcal{K}_1 = (\emptyset, \mathcal{D})$ where

$$\begin{aligned}\mathcal{D} \quad = \quad & \{\text{citizen\_of}(\text{Peter}, \text{UK}), \text{citizen\_of}(\text{Piotr}, \text{Poland}), \\ & \text{citizen\_of}(\text{Kazue}, \text{Japan}), \\ & \text{EuropC}(\text{UK}), \text{EuropC}(\text{Poland}), \\ & \text{Person}(\text{Peter}), \text{Person}(\text{Piotr})\}.\end{aligned}$$

Let $E^+ = \{\text{Peter}, \text{Piotr}\}$ and $E^- = \{\text{Kazue}\}$.

Task: Separate E using an $\mathcal{ALCI}$-concept.
Person separates E since

$$\mathcal{K}_1 \models \text{Person}(\text{Peter}) \quad \mathcal{K}_1 \models \text{Person}(\text{Piotr}) \quad \mathcal{K}_1 \not\models \text{Person}(\text{Kazue})$$

Examples

Let $\mathcal{K}_1 = (\emptyset, \mathcal{D})$ where

$$\begin{aligned} \mathcal{D} = \{ & \text{citizen\_of}(\text{Peter}, \text{UK}), \text{citizen\_of}(\text{Piotr}, \text{Poland}), \\ & \text{citizen\_of}(\text{Kazue}, \text{Japan}), \\ & \text{EuropC}(\text{UK}), \text{EuropC}(\text{Poland}), \\ & \text{Person}(\text{Peter}), \text{Person}(\text{Piotr})\}. \end{aligned}$$

Let $E^+ = \{\text{Peter}, \text{Piotr}\}$ and $E^- = \{\text{Kazue}\}$.

Task: Separate E using an $\mathcal{ALCI}$-concept.
Person separates E since

$$\mathcal{K}_1 \models \text{Person}(\text{Peter}) \quad \mathcal{K}_1 \models \text{Person}(\text{Piotr}) \quad \mathcal{K}_1 \not\models \text{Person}(\text{Kazue})$$

Not intended as only due to incompleteness.

Examples

Let $\mathcal{K}_1 = (\emptyset, \mathcal{D})$ where

$$\begin{aligned} \mathcal{D} \quad = \quad & \{\text{citizen\_of}(\text{Peter}, \text{UK}), \text{citizen\_of}(\text{Piotr}, \text{Poland}), \\ & \text{citizen\_of}(\text{Kazue}, \text{Japan}), \\ & \text{EuropC}(\text{UK}), \text{EuropC}(\text{Poland}), \\ & \text{Person}(\text{Peter}), \text{Person}(\text{Piotr})\}. \end{aligned}$$

Let $E^+ = \{\text{Peter}, \text{Piotr}\}$ and $E^- = \{\text{Kazue}\}$.

Task: Separate E using an $\mathcal{ALCI}$-concept.
Person separates E since

$\mathcal{K}_1 \models \text{Person}(\text{Peter}) \quad \mathcal{K}_1 \models \text{Person}(\text{Piotr}) \quad \mathcal{K}_1 \not\models \text{Person}(\text{Kazue})$

Not intended as only due to incompleteness.
Let

$$\mathcal{O} = \{\exists \text{citizen\_of}.\top \sqsubseteq \text{Person}\}$$

and $\mathcal{K}_2 = (\mathcal{O}, \mathcal{D})$. Then Person no longer separates.

Examples

Let $\mathcal{K}_1 = (\emptyset, \mathcal{D})$ where

$$\begin{aligned}\mathcal{D} \; = \; & \{\text{citizen\_of}(\text{Peter}, \text{UK}), \text{citizen\_of}(\text{Piotr}, \text{Poland}), \\ & \text{citizen\_of}(\text{Kazue}, \text{Japan}), \\ & \text{EuropC}(\text{UK}), \text{EuropC}(\text{Poland}), \\ & \text{Person}(\text{Peter}), \text{Person}(\text{Piotr})\}.\end{aligned}$$

Let $E^+ = \{\text{Peter}, \text{Piotr}\}$ and $E^- = \{\text{Kazue}\}$.

Task: Separate E using an $\mathcal{ALCI}$-concept.
Person separates E since

$$\mathcal{K}_1 \models \text{Person}(\text{Peter}) \quad \mathcal{K}_1 \models \text{Person}(\text{Piotr}) \quad \mathcal{K}_1 \not\models \text{Person}(\text{Kazue})$$

Not intended as only due to incompleteness.
Let

$$\mathcal{O} = \{\exists \text{citizen\_of}.\top \sqsubseteq \text{Person}\}$$

and $\mathcal{K}_2 = (\mathcal{O}, \mathcal{D})$. Then Person no longer separates. The concept $\exists$citizen_of.EuropC still separates.

Dimensions of the Separability Problem

Input Knowledge base $\mathcal{K}$, $E = (E^+, E^-)$.

Dimensions of the Separability Problem

Input Knowledge base $\mathcal{K}$, $E = (E^{+}, E^{-})$.
Question Is there a concept C separating E?

Dimensions of the Separability Problem

Input Knowledge base $\mathcal{K}$, $E = (E^+, E^-)$.
Question Is there a concept C separating E?

1. **Weak vs strong separability**
   in both cases: $\mathcal{K} \models C(a)$, for all $a \in E^+$
   weak: $\mathcal{K} \not\models C(a)$, for all $a \in E^-$
   strong: $\mathcal{K} \models \neg C(a)$, for all $a \in E^-$

Dimensions of the Separability Problem

Input Knowledge base $\mathcal{K}$, $E = (E^+, E^-)$.
Question Is there a concept C separating E?

1. Weak vs strong separability
in both cases: $\mathcal{K} \models C(a)$, for all $a \in E^+$
weak: $\mathcal{K} \not\models C(a)$, for all $a \in E^-$
strong: $\mathcal{K} \models \neg C(a)$, for all $a \in E^-$

To ensure strong separability in the previous example add:

$$\exists \text{citizen\_of}.\{\text{Japan}\} \sqsubseteq \neg \exists \text{citizen\_of}.\text{EuropC}$$

Dimensions of the Separability Problem

Input Knowledge base $\mathcal{K}$, $E = (E^+, E^-)$.
Question Is there a concept C separating E?

1. **Weak vs strong separability**
   in both cases: $\mathcal{K} \models C(a)$, for all $a \in E^+$
   weak: $\mathcal{K} \not\models C(a)$, for all $a \in E^-$
   strong: $\mathcal{K} \models \neg C(a)$, for all $a \in E^-$

To ensure strong separability in the previous example add:

$$\exists \text{citizen\_of}.\{\text{Japan}\} \sqsubseteq \neg \exists \text{citizen\_of}.\text{EuropC}$$

2. **Language of the separating formula**
   $\mathcal{EL}$, $\mathcal{ALC}$, $\mathcal{ALCI}$, FO, (U)CQs, ...

Dimensions of the Separability Problem

Input Knowledge base $\mathcal{K}$, $E = (E^+, E^-)$.
Question Is there a concept C separating E?

1. **Weak vs strong separability**
   in both cases: $\mathcal{K} \models C(a)$, for all $a \in E^+$
   weak: $\mathcal{K} \not\models C(a)$, for all $a \in E^-$
   strong: $\mathcal{K} \models \neg C(a)$, for all $a \in E^-$

To ensure strong separability in the previous example add:

$$\exists \text{citizen\_of}.\{\text{Japan}\} \sqsubseteq \neg\exists \text{citizen\_of}.\text{EuropC}$$

2. **Language of the separating formula**
   $\mathcal{EL}$, $\mathcal{ALC}$, $\mathcal{ALCI}$, FO, (U)CQs, ...

3. **Additional signature restrictions on C**
   For instance: only admit geography terms.

Dimensions of the Separability Problem

Input Knowledge base $\mathcal{K}$, $E = (E^+, E^-)$.
Question Is there a concept C separating E?

1. Weak vs strong separability
in both cases: $\mathcal{K} \models C(a)$, for all $a \in E^+$
weak: $\mathcal{K} \not\models C(a)$, for all $a \in E^-$
strong: $\mathcal{K} \models \neg C(a)$, for all $a \in E^-$

To ensure strong separability in the previous example add:

$$\exists \text{citizen\_of}.\{\text{Japan}\} \sqsubseteq \neg \exists \text{citizen\_of}.\text{EuropC}$$

2. Language of the separating formula
$\mathcal{EL}$, $\mathcal{ALC}$, $\mathcal{ALCI}$, FO, (U)CQs, ...

3. Additional signature restrictions on C
For instance: only admit geography terms.

4. Projective vs non-projective separability
Can C use symbols that do not occur in $\mathcal{K}$?

Projective vs non-projective separability

Consider $\mathcal{K} = (\emptyset, \mathcal{D})$ with

$$\begin{aligned} \mathcal{D} \quad = \quad & \{\text{born\_in}(\text{Peter}, \text{France}), \text{citizen\_of}(\text{Peter}, \text{France}), \\ & \text{born\_in}(\text{Kazue}, \text{Japan}), \text{citizen\_of}(\text{Kazue}, \text{Italy})\} \end{aligned}$$

Can we weakly separate $E = (\{\text{Peter}\}, \{\text{Kazue}\})$?

Projective vs non-projective separability

Consider $\mathcal{K} = (\emptyset, \mathcal{D})$ with

$$\begin{aligned}\mathcal{D} \quad = \quad & \{\text{born\_in}(\text{Peter}, \text{France}), \text{citizen\_of}(\text{Peter}, \text{France}), \\ & \text{born\_in}(\text{Kazue}, \text{Japan}), \text{citizen\_of}(\text{Kazue}, \text{Italy})\}\end{aligned}$$

Can we weakly separate $E = (\{\text{Peter}\}, \{\text{Kazue}\})$?

The FO-formula

$$\exists y(\text{citizen\_of}(x, y) \wedge \text{born\_in}(x, y))$$

weakly separates, but does not correspond to any $\mathcal{ALCI}$-concept.

Projective vs non-projective separability

Consider $\mathcal{K} = (\emptyset, \mathcal{D})$ with

$$\begin{aligned}\mathcal{D} = \; & \{\text{born\_in}(\text{Peter}, \text{France}), \text{citizen\_of}(\text{Peter}, \text{France}), \\ & \text{born\_in}(\text{Kazue}, \text{Japan}), \text{citizen\_of}(\text{Kazue}, \text{Italy})\}\end{aligned}$$

Can we weakly separate $E = (\{\text{Peter}\}, \{\text{Kazue}\})$?

The FO-formula

$$\exists y(\text{citizen\_of}(x, y) \wedge \text{born\_in}(x, y))$$

weakly separates, but does not correspond to any $\mathcal{ALCI}$-concept.

But still the following $\mathcal{ALCI}$-concept separates:

$$\forall \text{citizen\_of}.\text{EuropC} \rightarrow \exists \text{born\_in}.\text{EuropC}$$

Note that EuropC does not occur in $\mathcal{K}$.

Intermezzo: Conjunctive Queries

A conjunctive query (CQ) $q(x)$ is an FO-formula constructed from atoms $A(y)$ and $R(y_1, y_2)$ using $\exists$ and $\wedge$. We assume one free variable (answer variable). A union of conjunctive queries (UCQ) is a disjunction of CQs.

Intermezzo: Conjunctive Queries

A conjunctive query (CQ) $q(x)$ is an FO-formula constructed from atoms $A(y)$ and $R(y_1, y_2)$ using $\exists$ and $\wedge$. We assume one free variable (answer variable). A union of conjunctive queries (UCQ) is a disjunction of CQs.

For any UCQ $q(x)$, $(\mathcal{O}, \mathcal{D}) \models q(a)$ is defined as in FO. Has been investigated extensively over the past 15 by the DL community.

Intermezzo: Conjunctive Queries

A conjunctive query (CQ) $q(x)$ is an FO-formula constructed from atoms $A(y)$ and $R(y_1, y_2)$ using $\exists$ and $\wedge$. We assume one free variable (answer variable). A union of conjunctive queries (UCQ) is a disjunction of CQs.

For any UCQ $q(x)$, $(\mathcal{O}, \mathcal{D}) \models q(a)$ is defined as in FO. Has been investigated extensively over the past 15 by the DL community.

Associate with any $\mathcal{D}, a$ a CQ

$$\varphi_{\mathcal{D},a}$$

obtained by replacing individuals by variables and existentially quantifying over all variables distinct from the variables x replacing a.

Logically strongest CQ $\varphi(x)$ with $\mathcal{D} \models \varphi(a)$.

Example

For

$$\mathcal{D} = \{\text{citizen\_of}(\text{Peter}, \text{UK}), \text{citizen\_of}(\text{Piotr}, \text{Poland}), \text{citizen\_of}(\text{Kazue}, \text{Japan}), \text{EuropC}(\text{UK}), \text{EuropC}(\text{Poland}), \text{Person}(\text{Peter}), \text{Person}(\text{Piotr})\}$$

we have

$$\varphi_{\mathcal{D},\text{Peter}}(x) = \exists \vec{y}\ \text{citizen\_of}(x, y_{\text{UK}}) \wedge \text{citizen\_of}(y_{\text{Piotr}}, y_{\text{Poland}}) \wedge \text{citizen\_of}(y_{\text{Kazue}}, y_{\text{Japan}}) \wedge \text{EuropC}(y_{\text{UK}}) \wedge \text{EuropC}(y_{\text{Poland}}) \wedge \text{Person}(x) \wedge \text{Person}(y_{\text{Piotr}}).$$

Weak Separability for $\mathcal{ALCI}$

Assume $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ with $\mathcal{O}$ in $\mathcal{ALCI}$ and $E = (E^+, E^-)$ are given.
Then the following conditions are equivalent:

- E is projectively $\mathcal{ALCI}$-separable;
- E is FO-separable;
- E is UCQ-separable;

Weak Separability for $\mathcal{ALCI}$

Assume $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ with $\mathcal{O}$ in $\mathcal{ALCI}$ and $E = (E^+, E^-)$ are given.
Then the following conditions are equivalent:

- E is projectively $\mathcal{ALCI}$-separable;
- E is FO-separable;
- E is UCQ-separable;

Weak Separability for $\mathcal{ALCI}$

Assume $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ with $\mathcal{O}$ in $\mathcal{ALCI}$ and $E = (E^+, E^-)$ are given.
Then the following conditions are equivalent:

- E is projectively $\mathcal{ALCI}$-separable;
- E is FO-separable;
- E is UCQ-separable;
- the UCQ
$$\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$$
separates E.

Weak Separability for $\mathcal{ALCI}$

Assume $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ with $\mathcal{O}$ in $\mathcal{ALCI}$ and $E = (E^+, E^-)$ are given.
Then the following conditions are equivalent:

- E is projectively $\mathcal{ALCI}$-separable;
- E is FO-separable;
- E is UCQ-separable;
- the UCQ
$$\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$$
separates E.
- $\mathcal{K} \not\models \bigvee_{a \in E^+} \varphi_{\mathcal{D},a}(b)$ for $b \in E^-$.

Weak Separability for $\mathcal{ALCI}$

Assume $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ with $\mathcal{O}$ in $\mathcal{ALCI}$ and $E = (E^+, E^-)$ are given.
Then the following conditions are equivalent:

- E is projectively $\mathcal{ALCI}$-separable;
- E is FO-separable;
- E is UCQ-separable;
- the UCQ
$$\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$$
separates E.
- $\mathcal{K} \not\models \bigvee_{a \in E^+} \varphi_{\mathcal{D},a}(b)$ for $b \in E^-$.

Weak projective separability is polynomial time equivalent to the complement of rooted UCQ evaluation relative to $\mathcal{ALCI}$ KBs.

The latter problem is coNEXPTIME-complete (*Lutz 2008*).

Weak Separability: further discussion

- Point 5 implies: Projective $\mathcal{ALCI}$-separability is anti-monotone w.r.t. strengthening the ontology: if $\mathcal{O} \subseteq \mathcal{O}'$ and separability holds for $(\mathcal{O}', \mathcal{D})$, then it holds for $(\mathcal{O}, \mathcal{D})$.
- Slightly weaker version of the result holds for $\mathcal{ALC}$.

Weak Separability: further discussion

- Point 5 implies: Projective $\mathcal{ALCI}$-separability is anti-monotone w.r.t. strengthening the ontology: if $\mathcal{O} \subseteq \mathcal{O}'$ and separability holds for $(\mathcal{O}', \mathcal{D})$, then it holds for $(\mathcal{O}, \mathcal{D})$.
- Slightly weaker version of the result holds for $\mathcal{ALC}$.
- Non-projective $\mathcal{ALCI}$-separability: harder to analyse as it is very "syntax dependent". Still NExpTime-complete.

Strong Separability for $\mathcal{ALCI}$

For strong separability, additional symbols do not make a difference. So we do not distinguish between projective and non-projective strong separability.

Strong Separability for $\mathcal{ALCI}$

For strong separability, additional symbols do not make a difference. So we do not distinguish between projective and non-projective strong separability.

Assume $\mathcal{K} = (\mathcal{O}, \mathcal{D})$, $E = (E^+, E^-)$, with $\mathcal{O}$ an $\mathcal{ALCI}$- ontology. Then the following conditions are equivalent:

- E is strongly $\mathcal{ALCI}$-separable;
- E is strongly FO-separable;
- E is strongly UCQ-separable;
- the UCQ $\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$ strongly separates E;

Strong Separability for $\mathcal{ALCI}$

For strong separability, additional symbols do not make a difference. So we do not distinguish between projective and non-projective strong separability.

Assume $\mathcal{K} = (\mathcal{O}, \mathcal{D})$, $E = (E^+, E^-)$, with $\mathcal{O}$ an $\mathcal{ALCI}$- ontology. Then the following conditions are equivalent:

- E is strongly $\mathcal{ALCI}$-separable;
- E is strongly FO-separable;
- E is strongly UCQ-separable;
- the UCQ $\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$ strongly separates E;
- for all $a \in E^+$ and $b \in E^-$, the KB

  $$(\mathcal{O}, \mathcal{D}_{a=b}\mathcal{D})$$

  is unsatisfiable.

$\mathcal{D}_{a=b}\mathcal{D}$ obtained from $\mathcal{D}$ by adding a copy of $\mathcal{D}$ and then identifying a and b'.

Strong Separability for $\mathcal{ALCI}$

An $\mathcal{O}$-type is a maximal subset of the set of subconcepts of $\mathcal{O}$ that is satisfiable. There are exponentially many $\mathcal{O}$-types.

Strong Separability for $\mathcal{ALCI}$

An $\mathcal{O}$-type is a maximal subset of the set of subconcepts of $\mathcal{O}$ that is satisfiable. There are exponentially many $\mathcal{O}$-types.

Theorem For $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ an $\mathcal{ALCI}$-KB, the following conditions are equivalent:

- E is strongly $\mathcal{ALCI}$-separable;

Strong Separability for $\mathcal{ALCI}$

An $\mathcal{O}$-type is a maximal subset of the set of subconcepts of $\mathcal{O}$ that is satisfiable. There are exponentially many $\mathcal{O}$-types.

Theorem For $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ an $\mathcal{ALCI}$-KB, the following conditions are equivalent:

- E is strongly $\mathcal{ALCI}$-separable;
- For all $a \in E^+$ and $b \in E^-$, there do not exist models $\mathcal{I}$ and $\mathcal{J}$ of $\mathcal{K}$ such that $a^\mathcal{I}$ and $b^\mathcal{J}$ realize the same $\mathcal{O}$-type;

Strong Separability for $\mathcal{ALCI}$

An $\mathcal{O}$-type is a maximal subset of the set of subconcepts of $\mathcal{O}$ that is satisfiable. There are exponentially many $\mathcal{O}$-types.

Theorem For $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ an $\mathcal{ALCI}$-KB, the following conditions are equivalent:

- E is strongly $\mathcal{ALCI}$-separable;
- For all $a \in E^+$ and $b \in E^-$, there do not exist models $\mathcal{I}$ and $\mathcal{J}$ of $\mathcal{K}$ such that $a^\mathcal{I}$ and $b^\mathcal{J}$ realize the same $\mathcal{O}$-type;
- The $\mathcal{ALCI}$-concept $t_1 \sqcup \cdots \sqcup t_n$ strongly separates E, $t_1, \ldots, t_n$ the $\mathcal{O}$-types realizable in some $\mathcal{K}, a$, $a \in E^-$.

Strong Separability for $\mathcal{ALCI}$

An $\mathcal{O}$-type is a maximal subset of the set of subconcepts of $\mathcal{O}$ that is satisfiable. There are exponentially many $\mathcal{O}$-types.

Theorem For $\mathcal{K} = (\mathcal{O}, \mathcal{D})$ an $\mathcal{ALCI}$-KB, the following conditions are equivalent:

- E is strongly $\mathcal{ALCI}$-separable;
- For all $a \in E^+$ and $b \in E^-$, there do not exist models $\mathcal{I}$ and $\mathcal{J}$ of $\mathcal{K}$ such that $a^\mathcal{I}$ and $b^\mathcal{J}$ realize the same $\mathcal{O}$-type;
- The $\mathcal{ALCI}$-concept $t_1 \sqcup \cdots \sqcup t_n$ strongly separates E, $t_1, \ldots, t_n$ the $\mathcal{O}$-types realizable in some $\mathcal{K}$, a, $a \in E^-$.

Strong separability is ExpTime-complete

Strong Separability: further discussion

- A variation of the above works for $\mathcal{ALC}$.

Strong Separability: further discussion

- A variation of the above works for $\mathcal{ALC}$.
- Strong $\mathcal{ALCI}$-separability is coNP-complete in data complexity.

Impact of Signature Restrictions

None of the equivalences regarding separating power holds. For instance,

- as none of the symbols in $\mathcal{D}$ is necessarily in Σ, $\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$ does not work;
- anti-monotonicity does not hold as new axioms might introduce new conept names used in separatinf signature.

Impact of Signature Restrictions

None of the equivalences regarding separating power holds. For instance,

- as none of the symbols in $\mathcal{D}$ is necessarily in Σ, $\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$ does not work;
- anti-monotonicity does not hold as new axioms might introduce new conept names used in separatinf signature.

	without Σ	with Σ
weak projective $\mathcal{ALCI}$	NExpTime-c	2ExpTime-c
strong $\mathcal{ALCI}$	ExpTime-c	2ExpTime-c

Impact of Signature Restrictions

None of the equivalences regarding separating power holds. For instance,

- as none of the symbols in $\mathcal{D}$ is necessarily in Σ, $\bigvee_{a \in E^+} \varphi_{\mathcal{D},a}$ does not work;
- anti-monotonicity does not hold as new axioms might introduce new conept names used in separatinf signature.

	without Σ	with Σ
weak projective $\mathcal{ALCI}$	NExpTime-c	2ExpTime-c
strong $\mathcal{ALCI}$	ExpTime-c	2ExpTime-c

	without Σ	with Σ
projective weak	rooted UCQ answering	conservative extensions
strong	satisfiability	interpolant existence

Weak Separability and Conservative Extensions

$\mathcal{O} \cup \mathcal{O}^{\text{new}}$ is a conservative extension of $\mathcal{O}$ if

$$\mathcal{O} \cup \mathcal{O}^{\text{new}} \models C \sqsubseteq D \quad \Rightarrow \quad \mathcal{O} \models C \sqsubseteq D$$

for $\text{sig}(C \sqsubseteq D) \subseteq \text{sig}(\mathcal{O})$.

Weak Separability and Conservative Extensions

$\mathcal{O} \cup \mathcal{O}^{\text{new}}$ is a conservative extension of $\mathcal{O}$ if

$$\mathcal{O} \cup \mathcal{O}^{\text{new}} \models C \sqsubseteq D \quad \Rightarrow \quad \mathcal{O} \models C \sqsubseteq D$$

for $\text{sig}(C \sqsubseteq D) \subseteq \text{sig}(\mathcal{O})$.

Theorem [Lutz and W. 2006] Conservative Extensions in $\mathcal{ALCI}$ is 2ExpTime-complete.

Weak Separability and Conservative Extensions

$\mathcal{O} \cup \mathcal{O}^{\text{new}}$ is a conservative extension of $\mathcal{O}$ if

$$\mathcal{O} \cup \mathcal{O}^{\text{new}} \models C \sqsubseteq D \quad \Rightarrow \quad \mathcal{O} \models C \sqsubseteq D$$

for $\text{sig}(C \sqsubseteq D) \subseteq \text{sig}(\mathcal{O})$.

Theorem [Lutz and W. 2006] Conservative Extensions in $\mathcal{ALCI}$ is 2ExpTime-complete.

Reduction of conservative extensions to weak separability:

$C \sqsubseteq D$ is witness for non-conservativity
$\Leftrightarrow$ $\neg C \sqcup D$ separates $E = (\{a\}, \{b\})$ under $\mathcal{K} = (\mathcal{O}^*, \mathcal{D})$ and $\Sigma = \text{sig}(\mathcal{O})$ for

- $\mathcal{D} = \{A(a), B(b)\}$.
- $\mathcal{O}^* = (\mathcal{O} \cup \mathcal{O}^{\text{new}})^A \cup \mathcal{O}^B$, with $\mathcal{O}^C$ relativization of $\mathcal{O}$ to C.

Weak Separability and Conservative Extensions

$\mathcal{O} \cup \mathcal{O}^{\text{new}}$ is a conservative extension of $\mathcal{O}$ if

$$\mathcal{O} \cup \mathcal{O}^{\text{new}} \models C \sqsubseteq D \quad \Rightarrow \quad \mathcal{O} \models C \sqsubseteq D$$

for $\text{sig}(C \sqsubseteq D) \subseteq \text{sig}(\mathcal{O})$.

Theorem [Lutz and W. 2006] Conservative Extensions in $\mathcal{ALCI}$ is 2ExpTime-complete.

Reduction of conservative extensions to weak separability:

$C \sqsubseteq D$ is witness for non-conservativity
$\Leftrightarrow$ $\neg C \sqcup D$ separates $E = (\{a\}, \{b\})$ under $\mathcal{K} = (\mathcal{O}^*, \mathcal{D})$ and $\Sigma = \text{sig}(\mathcal{O})$ for

- $\mathcal{D} = \{A(a), B(b)\}$.
- $\mathcal{O}^* = (\mathcal{O} \cup \mathcal{O}^{\text{new}})^A \cup \mathcal{O}^B$, with $\mathcal{O}^C$ relativization of $\mathcal{O}$ to C.

Converse not obvious but proofs via emptiness for tree automata similar.

Strong Separability and Interpolant Existence

Assume $\mathcal{ALCIO}^u$ (nominals and the universal role u.)
A concept I is a Craig interpolant for $C \sqsubseteq D$ if

$$\text{sig}(I) \subseteq \text{sig}(C) \cap \text{sig}(D), \quad \models C \sqsubseteq I, \quad \models I \sqsubseteq D$$

Strong Separability and Interpolant Existence

Assume $\mathcal{ALCIO}^u$ (nominals and the universal role u.)
A concept I is a Craig interpolant for $C \sqsubseteq D$ if

$$\text{sig}(I) \subseteq \text{sig}(C) \cap \text{sig}(D), \quad \models C \sqsubseteq I, \quad \models I \sqsubseteq D$$

DLs with nominals do not have Craig interpolation, so the existence of I does not follow from $\models C \sqsubseteq D$.

Strong Separability and Interpolant Existence

Assume $\mathcal{ALCIO}^u$ (nominals and the universal role u.)
A concept I is a Craig interpolant for $C \sqsubseteq D$ if

$$\text{sig}(I) \subseteq \text{sig}(C) \cap \text{sig}(D), \quad \models C \sqsubseteq I, \quad \models I \sqsubseteq D$$

DLs with nominals do not have Craig interpolation, so the existence of I does not follow from $\models C \sqsubseteq D$.
Theorem [Artale et al 22] Interpolant existence is 2ExpTime complete for $\mathcal{ALCO}^u$ and $\mathcal{ALCIO}^u$.

Strong Separability and Interpolant Existence

Assume $\mathcal{ALCIO}^u$ (nominals and the universal role u.)
A concept I is a Craig interpolant for $C \sqsubseteq D$ if

$$\text{sig}(I) \subseteq \text{sig}(C) \cap \text{sig}(D), \quad \models C \sqsubseteq I, \quad \models I \sqsubseteq D$$

DLs with nominals do not have Craig interpolation, so the existence of I does not follow from $\models C \sqsubseteq D$.
Theorem [Artale et al 22] Interpolant existence is 2ExpTime complete for $\mathcal{ALCO}^u$ and $\mathcal{ALCIO}^u$.
Encode $\mathcal{K}, a$ and $\mathcal{K}, b$ in $\mathcal{ALCIO}^u$-concepts $C_{\Sigma,a}$ and $C_{\Sigma,b}$ sharing only Σ such that

$$
\begin{aligned}
& I \text{ strongly } \Sigma\text{-separates } \{a\} \text{ from } \{b\} \text{ in } \mathcal{K} \\
\Leftrightarrow \quad & \mathcal{K} \models I(a) \text{ and } \mathcal{K} \models \neg I(b) \\
\Leftrightarrow \quad & \models C_{\Sigma,a} \sqsubseteq I \text{ and } \models C_{\Sigma,b} \sqsubseteq \neg I \\
\Leftrightarrow \quad & I \text{ is Craig interpolant for } C_{\Sigma,a} \sqsubseteq \neg C_{\Sigma,b}
\end{aligned}
$$

Literature

- Jung, Lutz, Pulcini, Wolter: Logical separability of labeled data examples under ontologies. KR 2020 and AIJ 2022.
- Funk, Jung, Lutz, Pulcini, Wolter: Learning Description Logic Concepts: When can Positive and Negative Examples be Separated? IJCAI 2019
- Jung, Lutz, Pulcini, Wolter: Separating Positive and Negative Data Examples by Concepts and Formulas: The Case of Restricted Signatures. KR 2021
- Artale, Jung, Mazzullo, Ozaki, Wolter: Living Without Beth and Craig: Explicit Definitions and Interpolants in Description Logics with Nominals. AAAI 2021 and TOCL 2023.

Separability in Horn DLs

Introduction

Horn Description Logics are an important family of DLs in practice:
⇒ many real-world ontologies are (almost) Horn

Defining feature: Horn DLs are preserved under taking **products**

Basic members:

- $\mathcal{EL}$ $\quad C := \top \mid A \mid C \sqcap C \mid \exists r\,.\,C$

  $\exists\text{support}\,.\,\text{FootballClub}, \quad \text{Club} \sqcap \exists\text{competes\_in}\,.\,\text{League}$

- $\mathcal{ELI}$ $\quad C := \top \mid A \mid C \sqcap C \mid \exists R\,.\,C$ for a role name or an inverse role R

  $\exists\text{livesIn}\,.\,\text{NorthAmerica} \sqcap \exists\text{child}^-\,.\,\text{Rich}$

weak separability (strong separability meaningless without negation/⊥)

projective=non-projective

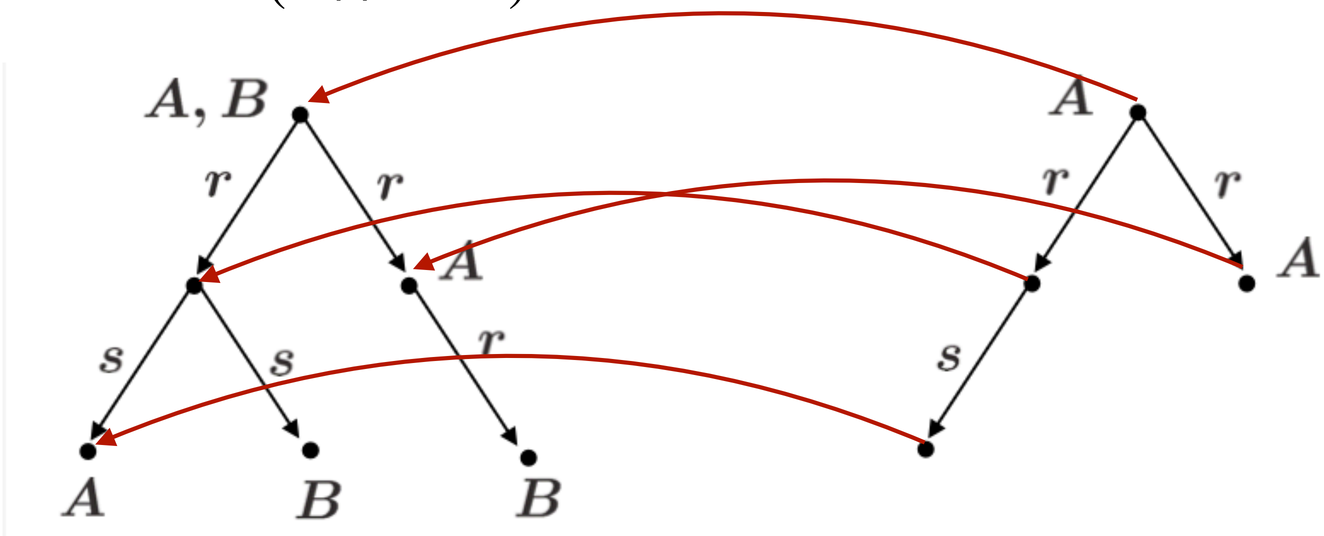
signature restrictions have no influence

Concepts ⇔ Databases

Examples

$C = A \sqcap B \sqcap$
$\exists r.(\exists s.A \sqcap \exists s.B) \sqcap$
$\exists r.(A \sqcap \exists r.B)$

$D = A \sqcap$
$\exists r.\exists s.\top \sqcap$
$\exists r.A$



Application query answering $\mathscr{D} \models C(a)$ = **Simulation** $(\mathscr{D}_C, a_C) \preceq (\mathscr{D}, a)$

Simulation $(\mathscr{D}, a) \preceq (\mathscr{D}', a')$ is a relation $S \subseteq \text{ind}(\mathscr{D}) \times \text{ind}(\mathscr{D}')$ with aSa' and

- bSb' and $B(b) \in \mathscr{D}$ implies $B(b') \in \mathscr{D}'$
- bSb' and $r(b, c) \in \mathscr{D}$ implies cSc' and $r(b', c') \in \mathscr{D}'$ for some c'

Simulation ≈ Homomorphism if $\mathscr{D}$ is tree-shaped

Direct Product

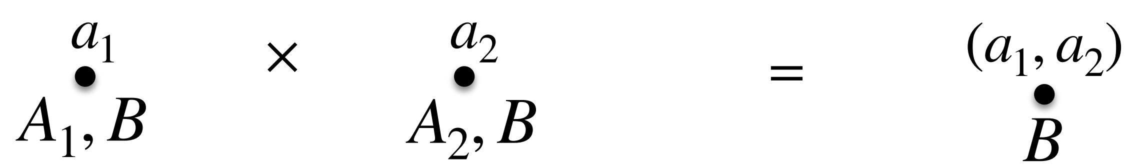
Direct Product $\mathcal{J} = \mathcal{I}_1 \times \mathcal{I}_2$ of interpretations $\mathcal{I}_1, \mathcal{I}_2$ is defined by

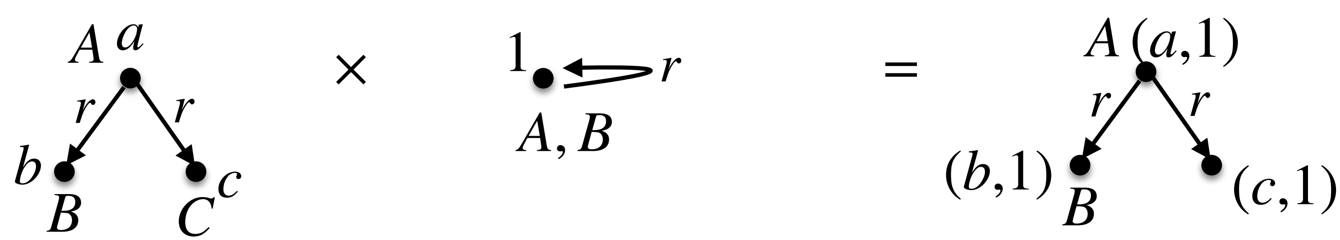
$$\Delta^{\mathcal{J}} = \Delta^{\mathcal{I}_1} \times \Delta^{\mathcal{I}_2}$$

$$A^{\mathcal{J}} = \{(d_1, d_2) \mid d_i \in A^{\mathcal{I}_i} \text{ für } i = 1{,}2\}$$

$$r^{\mathcal{J}} = \{((d_1, d_2), (e_1, e_2)) \mid (d_i, e_i) \in r^{\mathcal{I}_i} \text{ für } i = 1{,}2\}$$

Examples





Direct product of **databases** defined accordingly

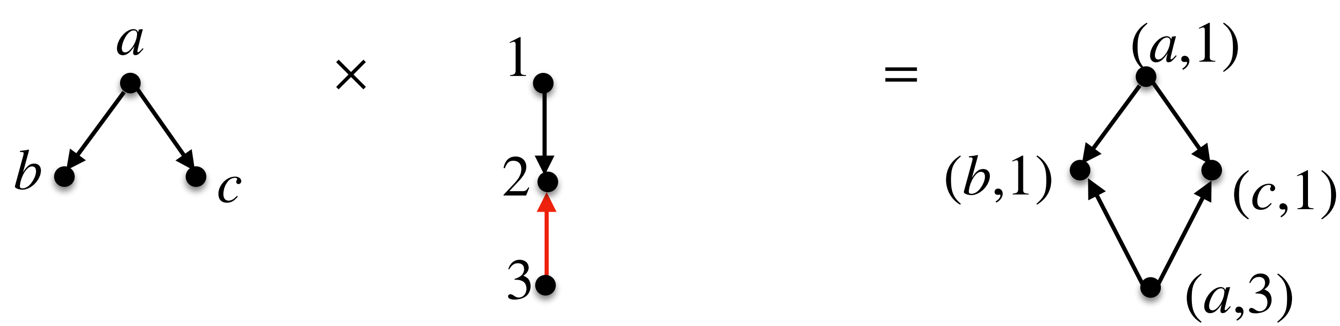
Properties of the Direct Product

For $\mathscr{EL}$ and $\mathscr{ELI}$-concepts C and databases $\mathscr{D}, \mathscr{D}'$ we have:

$$\mathscr{D} \times \mathscr{D}' \models C(a, a') \quad \Leftrightarrow \quad \mathscr{D} \models C(a) \text{ and } \mathscr{D}' \models C(a')$$

Product of any interpretation with $\mathscr{EL}$-concept results in $\mathscr{EL}$-concept

Not true for $\mathscr{ELI}$:



Product of n interpretations/databases with 2 domain elements has **size** 2^n**!**

Product Algorithm for $\mathscr{EL}$

[Baader et al, 90ies, Zarrieß & Turhan 2013, Barcelo & Romero 2017]

Characterization Let $E^+ = \{(\mathscr{D}_1, a_1), \ldots, (\mathscr{D}_n, a_n)\}, E^-$ be sets of examples. TFAE:

1. there is an $\mathscr{EL}$-concept separating E^+ and E^-

2. $(\Pi_i \mathscr{D}_i, (a_1, \ldots, a_n)) \not\preceq (\mathscr{D}, b)$ for all $(\mathscr{D}, b) \in E^-$ **(product simulation test)**

In case that all $\mathscr{D}_i$ „are“ $\mathscr{EL}$-concepts, we can compute the separating concept:

1. compute product $\mathscr{D}^* := \mathscr{D}_1 \times \ldots \times \mathscr{D}_n$
2. if $\mathscr{D}^*$ passes product simulation test for each $(\mathscr{D}, b) \in E^-$, return $\mathscr{D}^*$ (as a concept)
3. otherwise return „no separating concept“

Remarks

- if $\mathscr{D}_i$ are not $\mathscr{EL}$-concepts, we can still extract separating concept
- exponential time algorithm (product simulation test is NP-complete)
- characterization (with appropriate simulation) works for $\mathscr{ELI}$, but PSpace…ExpTime (however, product algorithm does **not** work!) [J, Lutz, Wolter 2019]

Extension 1: $\mathcal{EL}$-ontologies

Answering $\mathcal{EL}$-concept queries $C(a)$ under $\mathcal{EL}$-ontologies $\mathcal{O}$:

$$\mathcal{O}, \mathcal{D} \models C(a) \quad \text{if } a \in C^{\mathcal{I}} \text{ for all models } \mathcal{I} \text{ of } \mathcal{O} \text{ and } \mathcal{D}$$

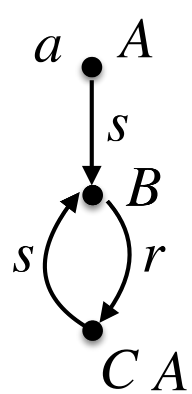
$\mathcal{EL}$-universal model

Given $\mathcal{O}, \mathcal{D}$, we can compute in polytime model $\mathcal{I}_{\mathcal{O},\mathcal{D}}$ of $\mathcal{O}, \mathcal{D}$ such that

$$\mathcal{O}, \mathcal{D} \models C(a) \text{ iff } \mathcal{I}_{\mathcal{O},\mathcal{D}} \models C(a) \text{ for all } \mathcal{EL}\text{-concepts } C$$

Chase-Like procedure:

For $\mathcal{O} = \{A \sqsubseteq \exists s\,.\,B, B \sqsubseteq \exists r\,.\,C, C \sqsubseteq A\}$ and $\mathcal{D} = \{A(a)\}$ we get:



Extension 1: $\mathscr{E}\mathscr{L}$-ontologies

$\mathscr{E}\mathscr{L}$-universal model

Given $\mathcal{O}, \mathscr{D}$, we can compute in polytime model $\mathscr{I}_{\mathcal{O},\mathscr{D}}$ of $\mathcal{O}, \mathscr{D}$ such that

$$\mathcal{O}, \mathscr{D} \models C(a) \text{ iff } \mathscr{I}_{\mathcal{O},\mathscr{D}} \models C(a) \text{ for all } \mathscr{E}\mathscr{L}\text{-concepts } C$$

Reduction of separability with ontologies to separability without ontologies

$$C \text{ separates } E^+ = \{(\mathscr{D}_1, a_1), \ldots, (\mathscr{D}_n, a_n)\} \text{ and } E^- = \{(\mathscr{E}_1, a_1), \ldots, (\mathscr{E}_k, a_k)\} \text{ under } \mathcal{O}$$

iff

$$C \text{ separates } \{(\mathscr{I}_{\mathcal{O},\mathscr{D}_1}, a_1), \ldots, (\mathscr{I}_{\mathcal{O},\mathscr{D}_n}, a_n)\} \text{ and } \{(\mathscr{I}_{\mathcal{O},\mathscr{E}_1}, a_1), \ldots, (\mathscr{I}_{\mathcal{O},\mathscr{E}_k}, a_k)\}$$

$\Rightarrow$ we can reuse product algorithm for $\mathscr{E}\mathscr{L}$

However

- complexity increases to ExpTime-complete [Funk 2019]
- size of smallest separating concept increases from poly to double exponential

Extension 2: $\mathcal{ELI}$-ontologies

$\mathcal{EL}$-universal model

Given $\mathcal{O}, \mathcal{D}$, we can compute in polytime model $\mathcal{I}_{\mathcal{O},\mathcal{D}}$ of $\mathcal{O}, \mathcal{D}$ such that

$$\mathcal{O}, \mathcal{D} \models C(a) \text{ iff } \mathcal{I}_{\mathcal{O},\mathcal{D}} \models C(a) \text{ for all } \mathcal{EL}\text{-concepts } C$$

$\mathcal{ELI}$-universal model

For every $\mathcal{O}, \mathcal{D}$, we can there is a model $\mathcal{I}_{\mathcal{O},\mathcal{D}}$ of $\mathcal{O}, \mathcal{D}$ such that

$$\mathcal{O}, \mathcal{D} \models C(a) \text{ iff } \mathcal{I}_{\mathcal{O},\mathcal{D}} \models C(a) \text{ for all } \mathcal{ELI}\text{-concepts } C$$

$\mathcal{ELI}$-universal model is infinite (and there is no finite one),
but **regular** and a representation can be computed in exponential time

Bad News regularity cannot be exploited: [Funk, J, Lutz, Pulcini, Wolter IJCAI 2019]
separability in $\mathcal{ELI}$ is **undecidable,** even for 2 positive + 1 negative example

(Notorious) open problem What about DL-Lite ontologies + $\mathcal{ELI}$-concept sep.?

Extension 3: Conjunctive Queries

We could also be interested in separability by **conjunctive queries (CQs)**

CQs generalize $\mathcal{EL}$/$\mathcal{ELI}$-concepts

Duality Conjunctive queries $\Leftrightarrow$ interpretations/databases

Answering conjunctive queries $q(x)$ under $\mathcal{EL}$-ontologies $\mathcal{O}$:

$$\mathcal{O}, \mathcal{D} \models q(a) \quad \text{if } (\mathcal{I}_q, x) \to (\mathcal{I}, a) \text{ for all models } \mathcal{I} \text{ of } \mathcal{O} \text{ and } \mathcal{D}$$

CQ-universal model for $\mathcal{EL}$-ontologies [Lutz, Toman, Wolter IJCAI 2009]

Given $\mathcal{O}, \mathcal{D}$, there is model $\mathcal{I}_{\mathcal{O},\mathcal{D}}$ of $\mathcal{O}, \mathcal{D}$ such that

$$\mathcal{O}, \mathcal{D} \models q(a) \text{ iff } \mathcal{I}_{\mathcal{O},\mathcal{D}} \models q(a) \text{ for all CQs } q(x)$$

$\mathcal{I}_{\mathcal{O},\mathcal{D}}$ is generally infinite, but
finite representation can be computed in polynomial time

Extension 3: Conjunctive Queries

[Gutierrez-Basulto, J, Sabellek IJCAI 2018]

Characterization Let $E^+ = \{(\mathscr{D}_1, a_1), \ldots, (\mathscr{D}_n, a_n)\}, E^-$ be sets of examples and $\mathcal{O}$ an $\mathscr{E}\mathscr{L}$-ontology. The following are equivalent:

1. there is an CQ E^+ and E^-
2. $(\Pi_i \mathscr{I}_{\mathcal{O},\mathscr{D}_i}, (a_1, \ldots, a_n)) \nrightarrow (\mathscr{I}_{\mathcal{O},\mathscr{D}}, b)$ for all $(\mathscr{D}, b) \in E^-$ (**product homomorphism test**)

Product homomorphism test is very similar to CQ separability without ontologies, which is coNExpTime-complete [Willard 2010, ten Cate & Dalmau 2015]

Product homomorphism test is decidable in coNExpTime exploiting the regularity [Gutierrez-Basulto, J, Sabellek IJCAI 2018]

Remarks

- separating CQs can be extracted, but double exponentially large (only exponential without ontologies)
- characterization (with appropriate universal model) works for $\mathscr{E}\mathscr{L}\mathscr{I}$, but CQ-separability under $\mathscr{E}\mathscr{L}\mathscr{I}$-ontologies also undecidable