TECHNISCHE
UNIVERSITÄT
DRESDEN

Faculty of Computer Science ● Institute of Theoretical Computer Science ● Chair of Automata Theory

OPTIMAL ABOX REPAIR W.R.T. STATIC $\mathcal{EL}$ TBOXES: FROM QUANTIFIED ABOXES BACK TO ABOXES

Franz Baader, Patrick Koopmann,
Francesco Kriegel, Adrian Nuradiansyah

Funded by the AI competence center ScaDS.AI Dresden/Leipzig
and by DFG in Project 430150274 and in TRR 248 (cpec, grant 389792660).

The 19th Extended Semantic Web Conference (ESWC 2022), 2 June 2022

Basics

- Description Logics (DLs): family of logic-based knowledge representation languages.
- OWL (the standard ontology language for the Semantic Web) is based on DLs.
- A DL ontology consists of an ABox (data) and a TBox (rules).

Basics

- Description Logics (DLs): family of logic-based knowledge representation languages.
- OWL (the standard ontology language for the Semantic Web) is based on DLs.
- A DL ontology consists of an ABox (data) and a TBox (rules).
- DL reasoners derive implicit consequences from ontologies.
- Unwanted consequences can be revealed, indicating errors in the explicitly represented axioms or privacy-sensitive information.
- Goal: repair the ontology to remove the unwanted consequences, preferably while retaining as many other consequences as possible.

Basics

- Description Logics (DLs): family of logic-based knowledge representation languages.
- OWL (the standard ontology language for the Semantic Web) is based on DLs.
- A DL ontology consists of an ABox (data) and a TBox (rules).
- DL reasoners derive implicit consequences from ontologies.
- Unwanted consequences can be revealed, indicating errors in the explicitly represented axioms or privacy-sensitive information.
- Goal: repair the ontology to remove the unwanted consequences, preferably while retaining as many other consequences as possible.
- Classical repair approaches based on axiom removal are not optimal, since information can get lost that does not contribute to the unwanted consequences.

The Description Logic $\mathcal{EL}$

- Concepts: $C \coloneqq \top \mid A \mid C \sqcap C \mid \exists r.C$
- Concept inclusions: $C \sqsubseteq D$
- Concept assertions: $C(a)$
- Role assertions: $r(a, b)$

Example: $\exists \mathit{parent}.(\mathit{Famous} \sqcap \mathit{Rich})$

Example: $\exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}$

Example: $(\exists \mathit{parent}.\mathit{Rich})(\mathit{BEN})$

Example: $\mathit{parent}(\mathit{BEN}, \mathit{JERRY})$

The Description Logic $\mathcal{EL}$

- Concepts: $C \coloneqq \top \mid A \mid C \sqcap C \mid \exists r.C$ Example: $\exists \mathit{parent}.(\mathit{Famous} \sqcap \mathit{Rich})$
- Concept inclusions: $C \sqsubseteq D$ Example: $\exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}$
- Concept assertions: $C(a)$ Example: $(\exists \mathit{parent}.\mathit{Rich})(\mathit{BEN})$
- Role assertions: $r(a, b)$ Example: $\mathit{parent}(\mathit{BEN}, \mathit{JERRY})$
- TBox: finite set of concept inclusions
- ABox: finite set of concept assertions and role assertions
- Polynomial-time reasoning

Previous Work

Franz Baader, Patrick Koopmann, Francesco Kriegel, Adrian Nuradiansyah:
Computing Optimal Repairs of Quantified ABoxes w.r.t. Static $\mathcal{EL}$ TBoxes. (CADE-28, 2021)

- ABox may contain errors, but the TBox is assumed to be correct.
- Instead of usual ABoxes, we considered an extended formalism.
- In addition to individuals, **quantified ABoxes (qABoxes)** $\exists X.\mathcal{A}$ can contain anonymous individuals (variables).

Previous Work

Franz Baader, Patrick Koopmann, Francesco Kriegel, Adrian Nuradiansyah:
Computing Optimal Repairs of Quantified ABoxes w.r.t. Static $\mathcal{EL}$ TBoxes. (CADE-28, 2021)

- ABox may contain errors, but the TBox is assumed to be correct.
- Instead of usual ABoxes, we considered an extended formalism.
- In addition to individuals, **quantified ABoxes (qABoxes)** $\exists X.\mathcal{A}$ can contain anonymous individuals (variables).
- Each ABox has an equivalent qABox.
  Example: $\{(A_1 \sqcap \exists r.A_2)(a), s(a,b)\}$ is equivalent to $\exists\{x\}.\{A_1(a), r(a,x), A_2(x), s(a,b)\}$

Previous Work

Franz Baader, Patrick Koopmann, Francesco Kriegel, Adrian Nuradiansyah:
Computing Optimal Repairs of Quantified ABoxes w.r.t. Static $\mathcal{EL}$ TBoxes. (CADE-28, 2021)

- ABox may contain errors, but the TBox is assumed to be correct.
- Instead of usual ABoxes, we considered an extended formalism.
- In addition to individuals, **quantified ABoxes (qABoxes)** $\exists X. \mathcal{A}$ can contain anonymous individuals (variables).
- Each ABox has an equivalent qABox.
  Example: $\{(A_1 \sqcap \exists r. A_2)(a), s(a,b)\}$ is equivalent to $\exists \{x\}. \{A_1(a), r(a,x), A_2(x), s(a,b)\}$
- Advantage: more consequences can be retained, optimal repairs exist.
- Disadvantage: currently no support for anonymous individuals by DL reasoners, although they are part of the OWL standard.

Previous Work

$\mathsf{Conj}(C_1 \sqcap C_2 \sqcap \cdots \sqcap C_n) = \{C_1, C_2, \ldots, C_n\}$

Repair Recipe: For each unwanted consequence $C(a)$ in the repair request $\mathcal{R}$:

- either choose a concept name $B \in \mathsf{Conj}(C)$ and remove $B(a)$ from $\mathcal{A}$,
- or choose an existential restriction $\exists r.\, D \in \mathsf{Conj}(C)$ and then, for each $r(a, b) \in \mathcal{A}$, either recursively modify $\mathcal{A}$ such that it does not entail $D(b)$ or remove $r(a, b)$ from $\mathcal{A}$.

Previous Work

$\mathsf{Conj}(C_1 \sqcap C_2 \sqcap \cdots \sqcap C_n) = \{C_1, C_2, \ldots, C_n\}$

Repair Recipe: For each unwanted consequence $C(a)$ in the repair request $\mathcal{R}$:

- either choose a concept name $B \in \mathsf{Conj}(C)$ and remove $B(a)$ from $\mathcal{A}$,
- or choose an existential restriction $\exists r.\, D \in \mathsf{Conj}(C)$ and then, for each $r(a, b) \in \mathcal{A}$, either recursively modify $\mathcal{A}$ such that it does not entail $D(b)$ or remove $r(a, b)$ from $\mathcal{A}$.

Forward Chaining:

- In order to not lose consequences that follow from removed axioms, but that do not itself violate the repair request, we initially **saturate** the qABox by means of the TBox.

Backward Chaining:

- When removing an atomic unwanted consequence $B(a)$ or $\exists r.\, D(a)$, it is also necessary to remove all $E(a)$ where $E \sqsubseteq^{\mathcal{T}} B$ or $E \sqsubseteq^{\mathcal{T}} \exists r.\, D$, respectively.
- It suffices to consider concepts $E \in \mathsf{Sub}(\mathcal{T}, \mathcal{R})$.

Previous Work

- **Canonical repairs** are based on the repair recipe, taking the TBox into account.
- To achieve **optimality**, we create enough copies of each object in the saturated input and modify each copy in another way.
  Example: For the unwanted consequence $(\exists \mathit{hasTopping}.(\mathit{Salami} \sqcap \mathit{Parmesan}))(\mathit{MYPIZZA})$, the qABox $\exists\{x\}.\{\mathit{hasTopping}(\mathit{MYPIZZA}, x), \mathit{Salami}(x), \mathit{Parmesan}(x)\}$ has the optimal repair $\exists\{x, y\}.\{\mathit{hasTopping}(\mathit{MYPIZZA}, x), \mathit{Salami}(x), \mathit{hasTopping}(\mathit{MYPIZZA}, y), \mathit{Parmesan}(y)\}$.
- We propose a rule-based approach to computing **optimized repairs**, which contain only relevant parts of the canonical repairs and are of exponential size only in the worst case.

Previous Work

- **Canonical repairs** are based on the repair recipe, taking the TBox into account.
- To achieve **optimality**, we create enough copies of each object in the saturated input and modify each copy in another way.
  Example: For the unwanted consequence $(\exists \mathit{hasTopping}.(\mathit{Salami} \sqcap \mathit{Parmesan}))(\mathit{MYPIZZA})$, the qABox $\exists\{x\}.\{\mathit{hasTopping}(\mathit{MYPIZZA}, x), \mathit{Salami}(x), \mathit{Parmesan}(x)\}$ has the optimal repair $\exists\{x, y\}.\{\mathit{hasTopping}(\mathit{MYPIZZA}, x), \mathit{Salami}(x), \mathit{hasTopping}(\mathit{MYPIZZA}, y), \mathit{Parmesan}(y)\}$.
- We propose a rule-based approach to computing **optimized repairs**, which contain only relevant parts of the canonical repairs and are of exponential size only in the worst case.
- Besides classical entailment between qABoxes (based on models), we considered entailment relations based on **instance queries (IQ)** and conjunctive queries (CQ).
- IQ-repairs retain as many instance queries as possible, i.e., consequences of the form $C(a)$.
- One main result: **All optimal IQ-repairs can be computed in exponential time, and each IQ-repair is entailed by an optimal one.**

Research Question in the Current Paper

Can we also obtain optimal repairs if the output of the repair process is restricted to usual ABoxes (without anonymous individuals)?

Example 1

- ABox: $\{parent(BEN, JERRY),\ Rich(JERRY)\}$
- TBox: $\{\exists parent.Rich \sqsubseteq Famous,\ Famous \sqsubseteq \exists friend.Famous,\ \exists friend.Famous \sqsubseteq Famous\}$
- Unwanted consequence: $Famous(BEN)$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\exists\{x, y\}.\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY}),\ \mathit{Famous}(\mathit{BEN}),\\ \mathit{friend}(\mathit{BEN}, x),\ \mathit{Famous}(x),\ \mathit{friend}(x, x),\\ \mathit{parent}(\mathit{BEN}, y),\ \mathit{Rich}(y)\}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\begin{aligned}\exists\{x,y\}.\{&\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY}),\ \mathit{Famous}(\mathit{BEN}),\\ &\mathit{friend}(\mathit{BEN}, x),\ \mathit{Famous}(x),\ \mathit{friend}(x, x),\\ &\mathit{parent}(\mathit{BEN}, y),\ \mathit{Rich}(y)\}\end{aligned}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$

- One optimal IQ-repair (obtained with our previous approach):

$$\exists\{x, y\}.\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY}),\ \mathit{Famous}(\mathit{BEN}),$$
$$\mathit{friend}(\mathit{BEN}, x),\ \mathit{Famous}(x),\ \mathit{friend}(x, x),$$
$$\mathit{parent}(\mathit{BEN}, y),\ \mathit{Rich}(y)\}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\exists\{x, y\}.\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY}),\ \mathit{Famous}(\mathit{BEN}),\\ \mathit{friend}(\mathit{BEN}, x),\ \mathit{Famous}(x),\ \mathit{friend}(x, x),\\ \mathit{parent}(\mathit{BEN}, y),\ \mathit{Rich}(y)\}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\begin{aligned}\exists\{x, y\}.\{&\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY}),\ \cancel{\mathit{Famous}(\mathit{BEN})},\\ &\mathit{friend}(\mathit{BEN}, x),\ \mathit{Famous}(x),\ \mathit{friend}(x, x),\\ &\mathit{parent}(\mathit{BEN}, y),\ \mathit{Rich}(y)\}\end{aligned}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN},\mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\begin{aligned}\exists\{x, y\}.\{&\xcancel{\mathit{parent}(\mathit{BEN},\mathit{JERRY})},\ \mathit{Rich}(\mathit{JERRY}),\ \xcancel{\mathit{Famous}(\mathit{BEN})},\\ &\mathit{friend}(\mathit{BEN}, x),\ \mathit{Famous}(x),\ \mathit{friend}(x, x),\\ &\mathit{parent}(\mathit{BEN}, y),\ \mathit{Rich}(y)\}\end{aligned}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\begin{aligned}\exists\{x, y\}.\{&\xcancel{\mathit{parent}(\mathit{BEN}, \mathit{JERRY})},\ \mathit{Rich}(\mathit{JERRY}),\ \xcancel{\mathit{Famous}(\mathit{BEN})},\\ &\mathit{friend}(\mathit{BEN}, x),\ \mathit{Famous}(x),\ \mathit{friend}(x, x),\\ &\mathit{parent}(\mathit{BEN}, y),\ \mathit{Rich}(y)\}\end{aligned}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN},\mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$

- One optimal IQ-repair (obtained with our previous approach):

$$\begin{aligned}\exists\{x,y\}.\{&\xcancel{\mathit{parent}(\mathit{BEN},\mathit{JERRY})},\ \mathit{Rich}(\mathit{JERRY}),\ \xcancel{\mathit{Famous}(\mathit{BEN})},\\ &\mathit{friend}(\mathit{BEN},x),\ \mathit{Famous}(x),\ \mathit{friend}(x,x),\\ &\mathit{parent}(\mathit{BEN},y),\ \xcancel{\mathit{Rich}(y)}\}\end{aligned}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN},\mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists\mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists\mathit{friend}.\mathit{Famous},\ \exists\mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\begin{aligned}\exists\{x,y\}.\{&\xcancel{\mathit{parent}(\mathit{BEN},\mathit{JERRY})},\ \mathit{Rich}(\mathit{JERRY}),\ \xcancel{\mathit{Famous}(\mathit{BEN})},\\ &\mathit{friend}(\mathit{BEN},x),\ \xcancel{\mathit{Famous}(x)},\ \mathit{friend}(x,x),\\ &\mathit{parent}(\mathit{BEN},y),\ \xcancel{\mathit{Rich}(y)}\}\end{aligned}$$

Example 1

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair (obtained with our previous approach):

$$\begin{aligned}\exists\{x, y\}.\{&\xcancel{\mathit{parent}(\mathit{BEN}, \mathit{JERRY})},\ \mathit{Rich}(\mathit{JERRY}),\ \xcancel{\mathit{Famous}(\mathit{BEN})},\\ &\mathit{friend}(\mathit{BEN}, x),\ \xcancel{\mathit{Famous}(x)},\ \mathit{friend}(x, x),\\ &\mathit{parent}(\mathit{BEN}, y),\ \xcancel{\mathit{Rich}(y)}\}\end{aligned}$$

- Due to the cycle $\mathit{friend}(x, x)$, this qABox entails $(\exists \mathit{friend}.)^n\top(\mathit{BEN})$ for each number $n \geq 0$.
- It is not equivalent to an ABox since these infinitely many consequences cannot be captured by an ABox, even with taking the TBox into account (the cycle is not covered by the TBox).
- In fact, there is no optimal ABox repair.

Example 2

- ABox: $\{\mathit{parent}(\mathit{BEN},\mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \xcancel{\exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$

Example 2

- ABox: $\{\mathit{parent}(\mathit{BEN}, \mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},\ \xcancel{\exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}}\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$
- One optimal IQ-repair:

$$\exists\{x, y\}.\{\mathit{parent}(\mathit{BEN}, x),\ \mathit{Rich}(\mathit{JERRY}),\ \mathit{friend}(\mathit{BEN}, y),\ \mathit{friend}(y, y),\ \mathit{Famous}(y)\}$$

Example 2

- ABox: $\{\mathit{parent}(\mathit{BEN},\mathit{JERRY}),\ \mathit{Rich}(\mathit{JERRY})\}$
- TBox: $\{\exists \mathit{parent}.\mathit{Rich} \sqsubseteq \mathit{Famous},\ \mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous},$ ~~$\exists \mathit{friend}.\mathit{Famous} \sqsubseteq \mathit{Famous}$~~$\}$
- Unwanted consequence: $\mathit{Famous}(\mathit{BEN})$

- One optimal IQ-repair:

$$\exists\{x,y\}.\{\mathit{parent}(\mathit{BEN},x),\ \mathit{Rich}(\mathit{JERRY}),\ \mathit{friend}(\mathit{BEN},y),\ \mathit{friend}(y,y),\ \mathit{Famous}(y)\}$$

- This qABox entails $(\exists \mathit{friend}.)^n\mathit{Famous}(\mathit{BEN})$ for each number $n \geq 0$.
- Since the TBox contains $\mathit{Famous} \sqsubseteq \exists \mathit{friend}.\mathit{Famous}$, all these consequences already follow from the assertion $(\exists \mathit{friend}.\mathit{Famous})(\mathit{BEN})$.
- Thus the repair is IQ-equivalent to the following ABox, taking the TBox into account:

$$\{(\exists \mathit{parent}.\top)(\mathit{BEN}),\ \mathit{Rich}(\mathit{JERRY}),\ (\exists \mathit{friend}.\mathit{Famous})(\mathit{BEN})\}$$

- The latter is an optimal ABox repair.

Definition of Optimal ABox Repairs

- **Repair request** $\mathcal{R}$: finite set of concept assertions
- An **ABox repair** of a qABox $\exists X.\mathcal{A}$ for $\mathcal{R}$ w.r.t. $\mathcal{T}$ is an ABox $\mathcal{B}$ such that
  - $\exists X.\mathcal{A} \models^{\mathcal{T}} \mathcal{B}$
  - $\mathcal{B} \not\models^{\mathcal{T}} C(a)$ for each $C(a) \in \mathcal{R}$
- $\mathcal{B}$ is **optimal** if there is no other ABox repair $\mathcal{C}$ such that $\mathcal{C} \models^{\mathcal{T}} \mathcal{B}$ but $\mathcal{B} \not\models^{\mathcal{T}} \mathcal{C}$.

Definition of Optimal ABox Repairs

- **Repair request** $\mathcal{R}$: finite set of concept assertions
- An **ABox repair** of a qABox $\exists X.\mathcal{A}$ for $\mathcal{R}$ w.r.t. $\mathcal{T}$ is an ABox $\mathcal{B}$ such that
  - $\exists X.\mathcal{A} \models^{\mathcal{T}} \mathcal{B}$
  - $\mathcal{B} \not\models^{\mathcal{T}} C(a)$ for each $C(a) \in \mathcal{R}$
- $\mathcal{B}$ is **optimal** if there is no other ABox repair $\mathcal{C}$ such that $\mathcal{C} \models^{\mathcal{T}} \mathcal{B}$ but $\mathcal{B} \not\models^{\mathcal{T}} \mathcal{C}$.

syntax independent

Definition of Optimal ABox Repairs

- **Repair request** $\mathcal{R}$: finite set of concept assertions
- An **ABox repair** of a qABox $\exists X.\mathcal{A}$ for $\mathcal{R}$ w.r.t. $\mathcal{T}$ is an ABox $\mathcal{B}$ such that
  - $\exists X.\mathcal{A} \models^{\mathcal{T}} \mathcal{B}$
  - $\mathcal{B} \not\models^{\mathcal{T}} C(a)$ for each $C(a) \in \mathcal{R}$
- $\mathcal{B}$ is **optimal** if there is no other ABox repair $\mathcal{C}$ such that $\mathcal{C} \models^{\mathcal{T}} \mathcal{B}$ but $\mathcal{B} \not\models^{\mathcal{T}} \mathcal{C}$.

syntax independent

- The previous examples show that optimal ABox repairs might not always exist.

IRQ-Entailment

- **IRQ-entailment** $\models^{\mathcal{T}}_{\mathsf{IRQ}}$ is slightly stronger than IQ-entailment $\models^{\mathcal{T}}_{\mathsf{IQ}}$ as it additionally takes role assertions into account.
- $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{IRQ}} \exists Y.\mathcal{B}$ if and only if $\exists Y.\mathcal{B} \models^{\mathcal{T}} \alpha$ implies $\exists X.\mathcal{A} \models^{\mathcal{T}} \alpha$ for each concept assertion $\alpha = C(a)$ or role assertion $\alpha = r(a, b)$.

IRQ-Entailment

- **IRQ-entailment** $\models^{\mathcal{T}}_{\mathsf{IRQ}}$ is slightly stronger than IQ-entailment $\models^{\mathcal{T}}_{\mathsf{IQ}}$ as it additionally takes role assertions into account.
- $\exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{IRQ}} \exists Y.\mathcal{B}$ if and only if $\exists Y.\mathcal{B} \models^{\mathcal{T}} \alpha$ implies $\exists X.\mathcal{A} \models^{\mathcal{T}} \alpha$ for each concept assertion $\alpha = C(a)$ or role assertion $\alpha = r(a, b)$.
- Classical entailment of an ABox by a qABox coincides with IRQ-entailment:

$$\exists X.\mathcal{A} \models^{\mathcal{T}} \mathcal{B} \text{ if and only if } \exists X.\mathcal{A} \models^{\mathcal{T}}_{\mathsf{IRQ}} \mathcal{B}$$

- Thus, **if we want to characterize the optimal ABox repairs of a (quantified) ABox w.r.t. classical entailment, we can investigate IRQ-repairs instead.**

Optimal IRQ-Repairs

- **(Optimal) IRQ-repairs** are defined like ABox repairs but with qABoxes instead of ABoxes and IRQ-entailment $\models^{\mathcal{T}}_{\mathsf{IRQ}}$ instead of classical entailment $\models^{\mathcal{T}}$.
- The set of canonical or optimized IQ-repairs (as in our CADE-28 paper) contains every optimal IRQ-repair (up to equivalence).
- Each IRQ-repair is entailed by an optimal one.
- **The set of all optimal IRQ-repairs can be computed in exponential time.**

Optimal IRQ-Repairs

- **(Optimal) IRQ-repairs** are defined like ABox repairs but with qABoxes instead of ABoxes and IRQ-entailment $\models^{\mathcal{T}}_{\mathsf{IRQ}}$ instead of classical entailment $\models^{\mathcal{T}}$.
- The set of canonical or optimized IQ-repairs (as in our CADE-28 paper) contains every optimal IRQ-repair (up to equivalence).
- Each IRQ-repair is entailed by an optimal one.
- **The set of all optimal IRQ-repairs can be computed in exponential time.**
- This is by chance since our definition of canonical IQ-repairs does not generate new role assertions between individuals and preserves as many of them as possible, although this is not necessary for IQ-entailment. Such a result might not hold for other sets of IQ-repairs.

Optimal IRQ-Repairs

- **(Optimal) IRQ-repairs** are defined like ABox repairs but with qABoxes instead of ABoxes and IRQ-entailment $\models^{\mathcal{T}}_{\mathsf{IRQ}}$ instead of classical entailment $\models^{\mathcal{T}}$.
- The set of canonical or optimized IQ-repairs (as in our CADE-28 paper) contains every optimal IRQ-repair (up to equivalence).
- Each IRQ-repair is entailed by an optimal one.
- **The set of all optimal IRQ-repairs can be computed in exponential time.**
- This is by chance since our definition of canonical IQ-repairs does not generate new role assertions between individuals and preserves as many of them as possible, although this is not necessary for IQ-entailment. Such a result might not hold for other sets of IQ-repairs.
- **Each optimal IRQ-repair is a representation of an optimal ABox repair, but we still need to transform it.**

Optimal ABox Approximations and Optimal ABox Repairs

- Given an optimal IRQ-repair, we try to find an **optimal ABox approximation**, which is an ABox that entails the same concept assertions and role assertions.
- If it exists, then it is an optimal ABox repair. Conversely, every optimal ABox repair can be obtained in this way.

Optimal ABox Approximations and Optimal ABox Repairs

- Given an optimal IRQ-repair, we try to find an **optimal ABox approximation**, which is an ABox that entails the same concept assertions and role assertions.
- If it exists, then it is an optimal ABox repair. Conversely, every optimal ABox repair can be obtained in this way.
- **Existence of an optimal ABox repair can be decided in exponential time, and all optimal ABox repairs can be computed in double-exponential time.**
- Contrary to the IQ- and IRQ-repairs, not every ABox repair is entailed by an optimal one.
- This corresponds to the optimal IRQ-repairs that do not have an optimal ABox approximation. These could still be transformed into an ABox by unfolding up to a fixed role depth. Increasing the role depth then always leads to better ABox repairs.

Computing Optimal ABox Approximations

- We first transform the given IRQ-repair into an equivalent **pre-approximation**.
- Existence and computation of an optimal ABox approximation of the pre-approximation is then reduced to the existence and computation of most specific concepts.

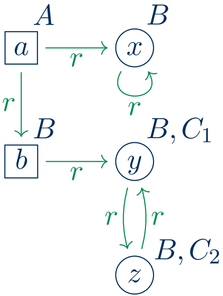
  Benjamin Zarrieß, Anni-Yasmin Turhan:
  Most specific generalizations w.r.t. general $\mathcal{EL}$-TBoxes. (IJCAI 2013)

Computing Optimal ABox Approximations

- We first transform the given IRQ-repair into an equivalent **pre-approximation**.
- Existence and computation of an optimal ABox approximation of the pre-approximation is then reduced to the existence and computation of most specific concepts.

  Benjamin Zarrieß, Anni-Yasmin Turhan:
  Most specific generalizations w.r.t. general $\mathcal{EL}$-TBoxes. (IJCAI 2013)
- **It can be decided in polynomial time if a qABox has an optimal ABox approximation. If so, the optimal ABox approximation can be computed in exponential time.**

Computing Optimal ABox Approximations



- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x), r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

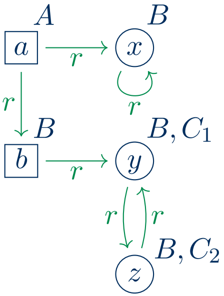
Computing Optimal ABox Approximations

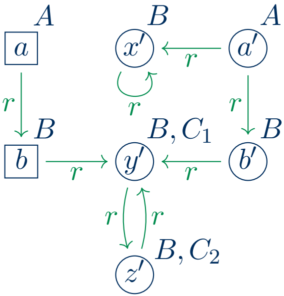
- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x),$$
$$r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed
  into the **pre-approximation** (see lower right):

$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$





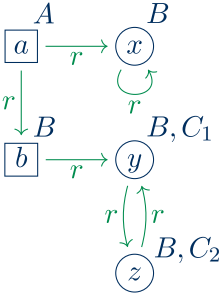
Computing Optimal ABox Approximations

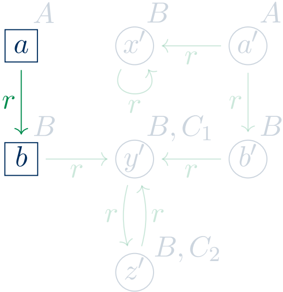
- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a),r(a,b),B(b),r(a,x),B(x),r(x,x),\\ r(b,y),B(y),C_1(y),r(y,z),B(z),C_2(z),r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\}\cup\mathcal{B}_a\cup\mathcal{B}_b$$





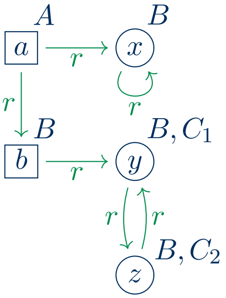
Computing Optimal ABox Approximations

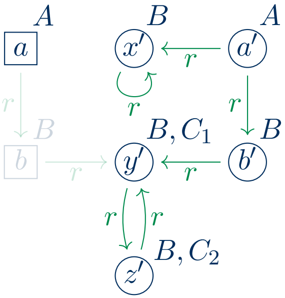
- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a),r(a,b),B(b),r(a,x),B(x),r(x,x),\\ r(b,y),B(y),C_1(y),r(y,z),B(z),C_2(z),r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\}\cup\mathcal{B}_a\cup\mathcal{B}_b$$





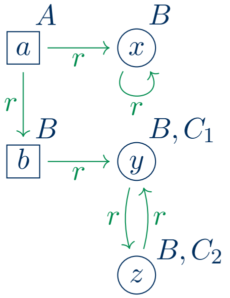
Computing Optimal ABox Approximations

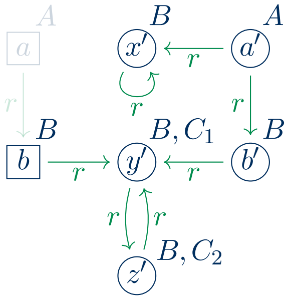
- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a),r(a,b),B(b),r(a,x),B(x),r(x,x),\\ r(b,y),B(y),C_1(y),r(y,z),B(z),C_2(z),r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\}\cup\mathcal{B}_a\cup\mathcal{B}_b$$





Computing Optimal ABox Approximations

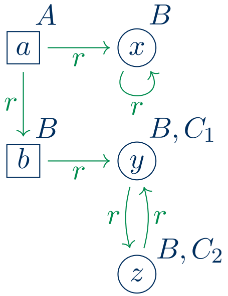
- Example: Saturated qABox (see upper right):

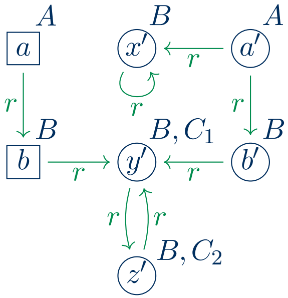
$$\exists\{x,y,z\}.\{A(a),r(a,b),B(b),r(a,x),B(x),r(x,x),$$
$$r(b,y),B(y),C_1(y),r(y,z),B(z),C_2(z),r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

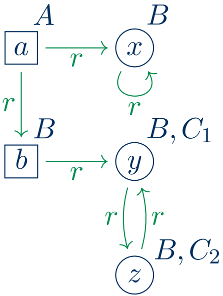
$$\{r(a,b)\}\cup\mathcal{B}_a\cup\mathcal{B}_b$$

- Optimal ABox approximation w.r.t. empty TBox:





Computing Optimal ABox Approximations



- Example: Saturated qABox (see upper right):

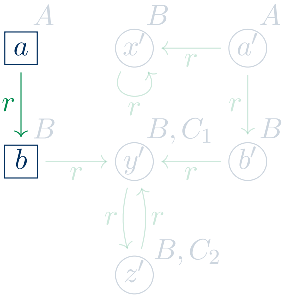
$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x),$$
$$r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

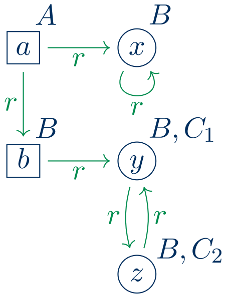
$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$

- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a,b)$$



Computing Optimal ABox Approximations



- Example: Saturated qABox (see upper right):

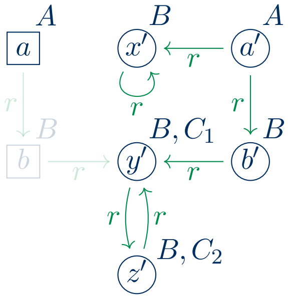
$$\exists\{x, y, z\}.\{A(a), r(a, b), B(b), r(a, x), B(x), r(x, x), r(b, y), B(y), C_1(y), r(y, z), B(z), C_2(z), r(z, y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

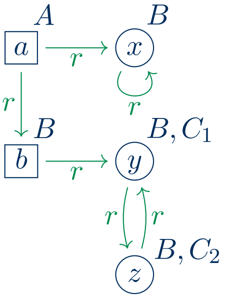
$$\{r(a, b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$

- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a, b), A(a)\}$$



Computing Optimal ABox Approximations

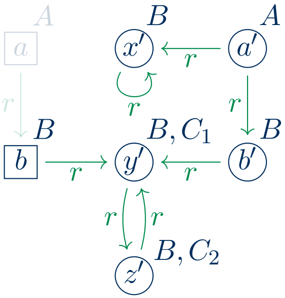


- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x), r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

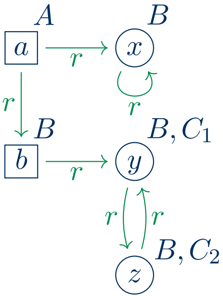
$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$



- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a,b), A(a), B(b)\}$$

Computing Optimal ABox Approximations



- Example: Saturated qABox (see upper right):

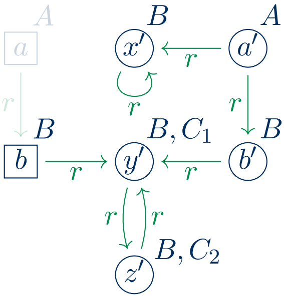
$$\exists\{x, y, z\}.\{A(a), r(a, b), B(b), r(a, x), B(x), r(x, x), r(b, y), B(y), C_1(y), r(y, z), B(z), C_2(z), r(z, y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

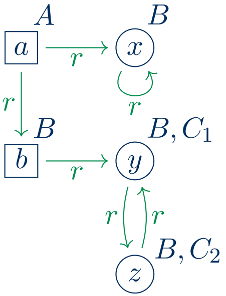
$$\{r(a, b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$

- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a, b), A(a), (B \sqcap \exists r.(B \sqcap C_1))(b)\}$$



Computing Optimal ABox Approximations



- Example: Saturated qABox (see upper right):

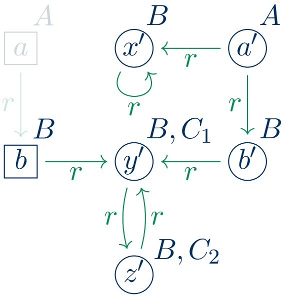
$$\exists\{x,y,z\}.\{A(a),r(a,b),B(b),r(a,x),B(x),r(x,x),\\ r(b,y),B(y),C_1(y),r(y,z),B(z),C_2(z),r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\}\cup\mathcal{B}_a\cup\mathcal{B}_b$$

- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a,b),A(a),(B\sqcap\exists r.(B\sqcap C_1\sqcap\exists r.(B\sqcap C_2)))(b)\}$$



Computing Optimal ABox Approximations

- Example: Saturated qABox (see upper right):

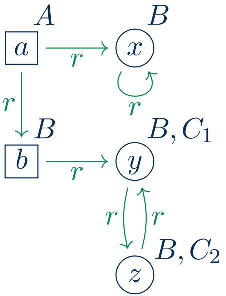
$$\exists\{x, y, z\}.\{A(a), r(a, b), B(b), r(a, x), B(x), r(x, x), \\ r(b, y), B(y), C_1(y), r(y, z), B(z), C_2(z), r(z, y)\}$$

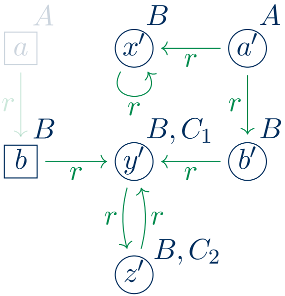
- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a, b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$

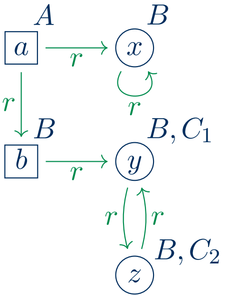
- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a, b), A(a), ???(b)\}$$





Computing Optimal ABox Approximations



- Example: Saturated qABox (see upper right):

$$\exists\{x, y, z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x), r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

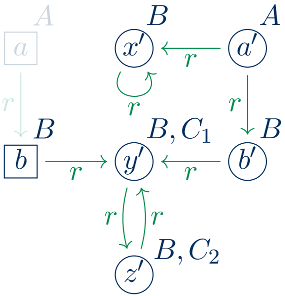
- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$

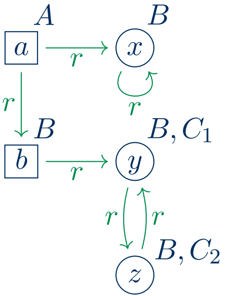
- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a,b), A(a), ???(b)\} \quad \lightning$$

There is no optimal ABox approximation.



Computing Optimal ABox Approximations

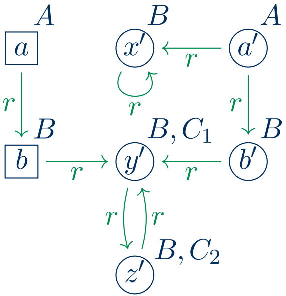


- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x), r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$



- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a,b), A(a), ???(b)\} \quad \lightning$$

There is no optimal ABox approximation.

- Optimal ABox approximation w.r.t. TBox $\{(B \sqcap C_1) \sqsubseteq \exists r.(B \sqcap C_2),\ (B \sqcap C_2) \sqsubseteq \exists r.(B \sqcap C_1)\}$:

Computing Optimal ABox Approximations

- Example: Saturated qABox (see upper right):

  $$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x), r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

  $$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$

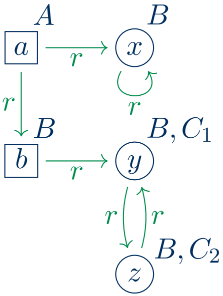
- Optimal ABox approximation w.r.t. empty TBox:

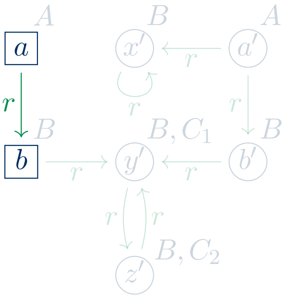
  $$\{r(a,b), A(a), ???(b)\} \quad ↯$$

  There is no optimal ABox approximation.

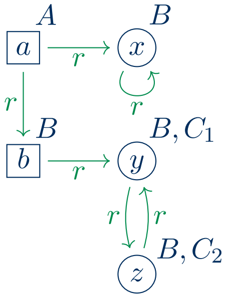
- Optimal ABox approximation w.r.t. TBox $\{(B \sqcap C_1) \sqsubseteq \exists r.(B \sqcap C_2),\ (B \sqcap C_2) \sqsubseteq \exists r.(B \sqcap C_1)\}$:

  $$\{r(a,b)$$





Computing Optimal ABox Approximations

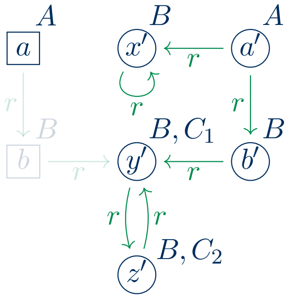


- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x),$$
$$r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$



- Optimal ABox approximation w.r.t. empty TBox:

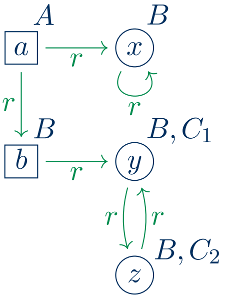
$\{r(a,b), A(a), ???(b)\}$ ↯

There is no optimal ABox approximation.

- Optimal ABox approximation w.r.t. TBox $\{(B \sqcap C_1) \sqsubseteq \exists r.(B \sqcap C_2),\ (B \sqcap C_2) \sqsubseteq \exists r.(B \sqcap C_1)\}$:

$\{r(a,b), A(a)\}$

Computing Optimal ABox Approximations

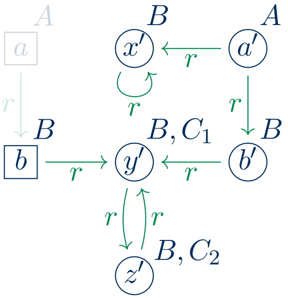


- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x), r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$



- Optimal ABox approximation w.r.t. empty TBox:

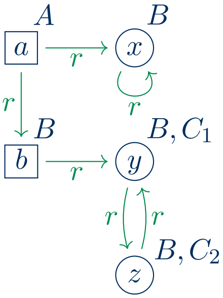
$$\{r(a,b), A(a), ???(b)\} \quad ↯$$

  There is no optimal ABox approximation.

- Optimal ABox approximation w.r.t. TBox $\{(B \sqcap C_1) \sqsubseteq \exists r.(B \sqcap C_2),\ (B \sqcap C_2) \sqsubseteq \exists r.(B \sqcap C_1)\}$:

$$\{r(a,b), A(a), B(b)\}$$

Computing Optimal ABox Approximations

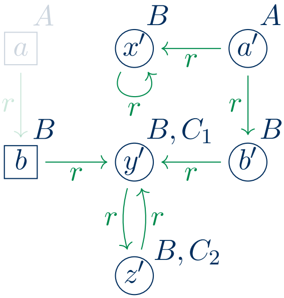


- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x), r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$

- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$



- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a,b), A(a), ???(b)\} \quad ↯$$

There is no optimal ABox approximation.

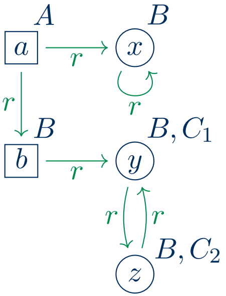
- Optimal ABox approximation w.r.t. TBox $\{(B \sqcap C_1) \sqsubseteq \exists r.(B \sqcap C_2),\ (B \sqcap C_2) \sqsubseteq \exists r.(B \sqcap C_1)\}$:

$$\{r(a,b), A(a), (B \sqcap \exists r.(B \sqcap C_1))(b)\}$$

Computing Optimal ABox Approximations

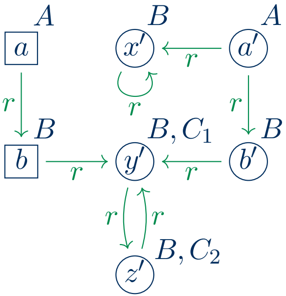
- Example: Saturated qABox (see upper right):

$$\exists\{x,y,z\}.\{A(a), r(a,b), B(b), r(a,x), B(x), r(x,x),$$
$$r(b,y), B(y), C_1(y), r(y,z), B(z), C_2(z), r(z,y)\}$$



- The saturated qABox is transformed into the **pre-approximation** (see lower right):

$$\{r(a,b)\} \cup \mathcal{B}_a \cup \mathcal{B}_b$$



- Optimal ABox approximation w.r.t. empty TBox:

$$\{r(a,b), A(a), ???(b)\} \quad ↯$$

There is no optimal ABox approximation.

- Optimal ABox approximation w.r.t. TBox $\{(B \sqcap C_1) \sqsubseteq \exists r.(B \sqcap C_2),\ (B \sqcap C_2) \sqsubseteq \exists r.(B \sqcap C_1)\}$:

$$\{r(a,b), A(a), (B \sqcap \exists r.(B \sqcap C_1))(b)\} \quad \checkmark$$

That’s it for now!

Do you have questions or comments?