TECHNISCHE UNIVERSITÄT DRESDEN

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

Abductive Differences of Quantified ABoxes

Francesco Kriegel

Theoretical Computer Science, Technische Universität Dresden, Germany
Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Germany

38th International Workshop on Description Logics, 3–6 September 2025

Quantified ABoxes

Quantified ABoxes

- A *quantified ABox* $\exists X.\mathcal{A}$ consists of
  - a finite set X of *variables* and
  - an $\mathcal{EL}$ ABox $\mathcal{A}$, called *matrix*, in which variables may be used in place of individuals.
- We assume every quantified ABox be in normal form, i.e. no complex concepts occur in the matrix.

Example. $\exists\emptyset.\{i : (A \sqcap \exists r.B),\ j : \top\}$ and $\exists\{x\}.\{i : A,\ (i,x) : r,\ x : B\}$ are equivalent, but only the latter is in formal form.

Quantified ABoxes

- A *quantified ABox* $\exists X.\mathcal{A}$ consists of
  - a finite set X of *variables* and
  - an $\mathcal{EL}$ ABox $\mathcal{A}$, called *matrix*, in which variables may be used in place of individuals.
- We assume every quantified ABox be in normal form,
  i.e. no complex concepts occur in the matrix.

Example. $\exists\emptyset.\{i : (A \sqcap \exists r.B),\ j : \top\}$ and $\exists\{x\}.\{i : A,\ (i,x) : r,\ x : B\}$ are equivalent, but only the latter is in formal form.

- The semantics is defined by models and variable assignments:
  $\mathcal{I} \models \exists X.\mathcal{A}$ iff. there is $\mathcal{Z} : X \to \mathrm{Dom}(\mathcal{I})$ such that $\mathcal{I}[\mathcal{Z}] \models \mathcal{A}$.

Quantified ABoxes

- Over signatures consisting of constants, unary predicates, and binary predicates only, the following are syntatic variants of each other, i.e. semantically the same:
  - relational structures with constants,
  - databases with nulls,
  - primitive-positive (pp) formulas in first-order logic,
  - conjunctive queries (CQs), and
  - quantified ABoxes.

Quantified ABoxes

- Over signatures consisting of constants, unary predicates, and binary predicates only, the following are syntatic variants of each other, i.e. semantically the same:
  - relational structures with constants,
  - databases with nulls,
  - primitive-positive (pp) formulas in first-order logic,
  - conjunctive queries (CQs), and
  - quantified ABoxes.
- We can thus reuse results for any of the above.
- Most importantly: $\exists X.\mathcal{A} \models \exists Y.\mathcal{B}$ iff. there is a homomorphism from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$.

Quantified ABoxes

- Over signatures consisting of constants, unary predicates, and binary predicates only, the following are syntatic variants of each other, i.e. semantically the same:
  - relational structures with constants,
  - databases with nulls,
  - primitive-positive (pp) formulas in first-order logic,
  - conjunctive queries (CQs), and
  - quantified ABoxes.
- We can thus reuse results for any of the above.
- Most importantly: $\exists X.\mathcal{A} \models \exists Y.\mathcal{B}$ iff. there is a homomorphism from $\exists Y.\mathcal{B}$ to $\exists X.\mathcal{A}$.

  Recall: a *homomorphism* from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B}$ is a mapping $h\colon \mathsf{Obj}(\exists X.\mathcal{A}) \to \mathsf{Obj}(\exists Y.\mathcal{B})$ that fulfills the following conditions:

  1. $h(i) = i$ for each individual i,
  2. if $t : A \in \mathcal{A}$, then $h(t) : A \in \mathcal{B}$,
  3. if $(t, u) : r \in \mathcal{A}$, then $(h(t), h(u)) : r \in \mathcal{B}$.

Explaining Observations by Abductive Differences

Observations and Explanations

Definition. Consider two quantified ABoxes:

- an observation $\exists X.\mathcal{A}$
- and a knowledge base $\exists Y.\mathcal{B}$.

An *abductive difference* (or *explanation*) of $\exists X.\mathcal{A}$ w.r.t. $\exists Y.\mathcal{B}$ is a quantified ABox $\exists Z.\mathcal{C}$ such that $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C} \models \exists X.\mathcal{A}$.

Observations and Explanations

Definition. Consider two quantified ABoxes:

- an observation $\exists X.\mathcal{A}$
- and a knowledge base $\exists Y.\mathcal{B}$.

An *abductive difference* (or *explanation*) of $\exists X.\mathcal{A}$ w.r.t. $\exists Y.\mathcal{B}$ is a quantified ABox $\exists Z.\mathcal{C}$ such that $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C} \models \exists X.\mathcal{A}$.

It is *minimal* if there is no other explanation $\exists Z'.\mathcal{C}'$ with $\exists Z.\mathcal{C} \models \exists Z'.\mathcal{C}'$ but $\exists Z'.\mathcal{C}' \not\models \exists Z.\mathcal{C}$.

Observations and Explanations

Definition. Consider two quantified ABoxes:

- an observation $\exists X.\mathcal{A}$
- and a knowledge base $\exists Y.\mathcal{B}$.

An *abductive difference* (or *explanation*) of $\exists X.\mathcal{A}$ w.r.t. $\exists Y.\mathcal{B}$ is a quantified ABox $\exists Z.\mathcal{C}$ such that $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C} \models \exists X.\mathcal{A}$.

It is *minimal* if there is no other explanation $\exists Z'.\mathcal{C}'$ with $\exists Z.\mathcal{C} \models \exists Z'.\mathcal{C}'$ but $\exists Z'.\mathcal{C}' \not\models \exists Z.\mathcal{C}$.

Example.

- Observation: $\exists\{x\}.\{\mathsf{tom} : \mathsf{Cat},\ (\mathsf{tom}, x) : \mathsf{chases},\ x : \mathsf{Mouse}\}$
- Knowledge base: $\exists\emptyset.\{\mathsf{tom} : \mathsf{Cat},\ \mathsf{jerry} : \mathsf{Mouse}\}$
- Two minimal explanations: $\exists\{x\}.\{(\mathsf{tom}, x) : \mathsf{chases},\ x : \mathsf{Mouse}\}$ and $\exists\emptyset.\{(\mathsf{tom}, \mathsf{jerry}) : \mathsf{chases}\}$

Observations and Explanations

Definition. Consider two quantified ABoxes:

- an observation $\exists X.\mathcal{A}$
- and a knowledge base $\exists Y.\mathcal{B}$.

An *abductive difference* (or *explanation*) of $\exists X.\mathcal{A}$ w.r.t. $\exists Y.\mathcal{B}$ is a quantified ABox $\exists Z.\mathcal{C}$ such that $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C} \models \exists X.\mathcal{A}$.

It is *minimal* if there is no other explanation $\exists Z'.\mathcal{C}'$ with $\exists Z.\mathcal{C} \models \exists Z'.\mathcal{C}'$ but $\exists Z'.\mathcal{C}' \not\models \exists Z.\mathcal{C}$.

How can we compute all minimal abductive differences?

Lower Bound

An observation can have at least exponentially many explanations.

Example. For each number $n \geq 1$, consider

- the observation $\exists\{x_1, \ldots, x_n\}.\{(x_1, x_2) : r,\ (x_2, x_3) : r,\ \ldots,\ (x_{n-1}, x_n) : r,\ x_1 : A_1,\ \ldots,\ x_n : A_n\}$
- and the knowledge base $\exists\emptyset.\{(i, i) : r,\ (j, j) : r,\ (i, j) : r,\ (j, i) : r\}$.

Then, in order to obtain a minimal explanation, we can choose between $i : A_\ell$ and $j : A_\ell$ for each $\ell \in \{1, \ldots, n\}$, i.e. every qABox $\exists\emptyset.\{t_1 : A_1,\ \ldots,\ t_n : A_n\}$ with $t_\ell \in \{i, j\}$ is a minimal explanation. Thus there are at least 2^n minimal explanations.

Partial Homomorphisms

Consider an observation $\exists X.\mathcal{A}$, a knowledge base $\exists Y.\mathcal{B}$, and an explanation $\exists Z.\mathcal{C}$.

Then $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C} \models \exists X.\mathcal{A}$
and thus there is a homomorphism h from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C}$.

Partial Homomorphisms

Consider an observation $\exists X.\mathcal{A}$, a knowledge base $\exists Y.\mathcal{B}$, and an explanation $\exists Z.\mathcal{C}$.

Then $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C} \models \exists X.\mathcal{A}$
and thus there is a homomorphism h from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C}$.

We split h into two mappings:

1. p is the part of h that maps to objects of the knowledge base $\exists Y.\mathcal{B}$,
2. q is the part of h that maps to objects of the explanation $\exists Z.\mathcal{C}$.

Partial Homomorphisms

Consider an observation $\exists X.\mathcal{A}$, a knowledge base $\exists Y.\mathcal{B}$, and an explanation $\exists Z.\mathcal{C}$.

Then $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C} \models \exists X.\mathcal{A}$
and thus there is a homomorphism h from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B} \cup \exists Z.\mathcal{C}$.

We split h into two mappings:

1. p is the part of h that maps to objects of the knowledge base $\exists Y.\mathcal{B}$,
2. q is the part of h that maps to objects of the explanation $\exists Z.\mathcal{C}$.

p is a partial function from $\mathsf{Obj}(\exists X.\mathcal{A})$ to $\mathsf{Obj}(\exists Y.\mathcal{B})$ that pinpoints the part of the observation that is already known.

We call p a *partial homomorphism* from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B}$ (see paper for details).
This notion is independent from the particular explanation $\exists Z.\mathcal{C}$ and the part q.

All minimal abductive differences can be obtained from these partial homomorphisms.

p-Differences

For each partial homomorphism p from the observation $\exists X.\mathcal{A}$ to the knowledge base $\exists Y.\mathcal{B}$, we can construct the *p-difference* $\exists X.\mathcal{A} \setminus^{p} \exists Y.\mathcal{B}$.

The p-difference consists of the observation part that is unknown according to p.

p-Differences

For each partial homomorphism p from the observation $\exists X.\mathcal{A}$ to the knowledge base $\exists Y.\mathcal{B}$, we can construct the *p-difference* $\exists X.\mathcal{A} \setminus^p \exists Y.\mathcal{B}$.

The p-difference consists of the observation part that is unknown according to p.

p-differences are canonical:

1. p can be extended to a homomorphism from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B} \cup (\exists X.\mathcal{A} \setminus^p \exists Y.\mathcal{B})$.
   Thus, every p-difference is an abductive difference.

p-Differences

For each partial homomorphism p from the observation $\exists X.\mathcal{A}$ to the knowledge base $\exists Y.\mathcal{B}$, we can construct the *p-difference* $\exists X.\mathcal{A} \setminus^p \exists Y.\mathcal{B}$.

The p-difference consists of the observation part that is unknown according to p.

p-differences are canonical:

1. p can be extended to a homomorphism from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B} \cup (\exists X.\mathcal{A} \setminus^p \exists Y.\mathcal{B})$.
   Thus, every p-difference is an abductive difference.
2. Each explanation entails some p-difference.
   Thus, every minimal explanation is equivalent to a p-difference.

Theorem. Up to equivalence, each minimal explanation has polynomial size and the set of all minimal explanations can be computed in exponential time.

Outlook

Implementation and Evaluation

There is a correspondence between

- partial homomorphisms from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B}$
- and homomorphisms from $\exists X.\mathcal{A}$ to an extension of $\exists Y.\mathcal{B}$.

Thus, partial homomorphisms can be enumerated with off-the-shelf query-answering systems.

Implementation and Evaluation

There is a correspondence between

- partial homomorphisms from $\exists X.\mathcal{A}$ to $\exists Y.\mathcal{B}$
- and homomorphisms from $\exists X.\mathcal{A}$ to an extension of $\exists Y.\mathcal{B}$.

Thus, partial homomorphisms can be enumerated with off-the-shelf query-answering systems.

Interesting future work:

- Implementation
- Evaluation with real-world datasets

Taking Ontologies into Account

(Minimal) abductive differences can also be considered w.r.t. ontologies.
An observation can then have infinitely many non-equivalent explanations,
and their sizes are not bounded.

Example. Consider

- the observation $\{\mathsf{alice} : \mathsf{Human}\}$
- and the KB consisting of the $\mathcal{EL}$ ABox $\{\mathsf{bob} : \mathsf{Human}\}$ and the $\mathcal{EL}$ ontology $\{\exists \mathsf{hasParent}.\mathsf{Human} \sqsubseteq \mathsf{Human}\}$.

For each number $n > 0$, the qABox $\exists\{x_1, \ldots, x_n\}.\{(\mathsf{alice}, x_1) : \mathsf{hasParent},\ (x_1, x_2) : \mathsf{hasParent},\ \ldots,\ (x_{n-1}, x_n) : \mathsf{hasParent},\ (x_n, \mathsf{bob}) : \mathsf{hasParent}\}$ is a minimal abductive difference.

Taking Ontologies into Account

(Minimal) abductive differences can also be considered w.r.t. ontologies.
An observation can then have infinitely many non-equivalent explanations, and their sizes are not bounded.

Example. Consider

- the observation $\{\mathsf{alice} : \mathsf{Human}\}$
- and the KB consisting of the $\mathcal{EL}$ ABox $\{\mathsf{bob} : \mathsf{Human}\}$ and the $\mathcal{EL}$ ontology $\{\exists\mathsf{hasParent}.\mathsf{Human} \sqsubseteq \mathsf{Human}\}$.

For each number $n > 0$, the qABox $\exists\{x_1, \dots, x_n\}.\{(\mathsf{alice}, x_1) : \mathsf{hasParent}, (x_1, x_2) : \mathsf{hasParent}, \dots, (x_{n-1}, x_n) : \mathsf{hasParent}, (x_n, \mathsf{bob}) : \mathsf{hasParent}\}$ is a minimal abductive difference.

Interesting future work:

- Enumeration of all minimal explanations
- Use of practically motivated metrics to restrict and compare explanations
- User interaction to pinpoint one practically useful explanation