Towards a Thorough Understanding of ABox Abduction Under Repair Semantics

Anselm Haak[1] Patrick Koopmann[2] Yasir Mahmood[3] Anni-Yasmin Turhan[1]

[1]Knowledge Representation Group, Paderborn University

[2]Knowledge in Artificial Intelligence, Vrije Universiteit Amsterdam

[3]Data Science Group, Paderborn University

July 24, 2026

ABox abduction

- ABox abduction explains missing entailments
- our setting:
  - knowledge base (KB) $\mathcal{K}$
  - observation: concept assertion (fact) α that should be entailed
  - if $\mathcal{K} \not\models \alpha$, why not?
  - explanation / hypothesis: set of ABox assertions (facts) $\mathcal{H}$ s.t. $\mathcal{K} \cup \mathcal{H} \models \alpha$

[Elsenbroich, Kutz, Sattler '06]

- What happens in the case of dirty data / inconsistency?

[Du, Wang, Shen '15]
[Bienvenu, Bourgaux, Goasdoué '19]
[Lukasiewicz, Malizia, Molinaro '22]

- systematic study: complexity under repair semantics for $\mathcal{EL}_\perp$ and DL-Lite

ABox abduction in $\mathcal{EL}_\bot$

- $\mathcal{EL}_\bot$ TBox (terminological part of KB):
  concept inclusions $C \sqsubseteq C$, with

$$C ::= C \sqcap C \mid \exists r.C \mid A \mid \top \mid \bot$$

  r role name, A concept name

- simple ABox:
  concept assertions $A(b)$, role assertions $r(a, b)$
  A concept name, r role name, a, b individuals
- observation: concept assertion $A(b)$
- hypothesis: set of concept and role assertions

Example: ABox abduction in medical diagnosis

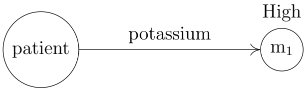
TBox $\mathcal{T}$:

$$\mathrm{High} \sqcap \mathrm{Low} \sqsubseteq \bot$$

$$\exists \mathrm{potassium}.\mathrm{Low} \sqcap \mathrm{CardiacDisease} \sqsubseteq \mathrm{CardiacArrest}$$
$$\exists \mathrm{potassium}.\mathrm{High} \sqcap \mathrm{CardiacDisease} \sqsubseteq \mathrm{CardiacArrest}$$

$$\exists \mathrm{potassium}.\mathrm{High} \sqcap \exists \mathrm{calcium}.\mathrm{Low} \sqsubseteq \mathrm{CardiacArrest}$$

ABox $\mathcal{A}$:



- observation: $\mathrm{CardiacArrest}(\mathrm{patient})$
- intended hypotheses:
  $\{\mathrm{CardiacDisease}(\mathrm{patient})\}$ and $\{\mathrm{Low}(\mathrm{m}_2), \mathrm{calcium}(\mathrm{patient}, \mathrm{m}_2)\}$
- trivial hypothesis: $\{\mathrm{CardiacArrest}(\mathrm{patient})\}$

Handling dirty data: repair semantics

- $\mathcal{K} = \mathcal{T} \cup \mathcal{A}$ inconsistent KB $\quad\rightsquigarrow\quad$ classical entailment useless
- conflict of $\mathcal{K}$: $\subseteq$-minimal $\mathcal{T}$-inconsistent subset of $\mathcal{A}$
- (ABox) repair of $\mathcal{K}$: $\subseteq$-maximal $\mathcal{T}$-consistent subset of $\mathcal{A}$

Definition (AR and Brave semantics)

Inconsistent KB $\mathcal{K} = \mathcal{T} \cup \mathcal{A}$, concept assertion α.

$\mathcal{T} \cup \mathcal{A} \models_{\mathrm{AR}} \alpha$ iff $\mathcal{T} \cup \mathcal{R} \models \alpha$ for all repairs $\mathcal{R}$ of $\mathcal{K}$

$\mathcal{T} \cup \mathcal{A} \models_{\mathrm{Brave}} \alpha$ iff $\mathcal{T} \cup \mathcal{R} \models \alpha$ for some repair $\mathcal{R}$ of $\mathcal{K}$

	$\mathcal{EL}_\bot$	DL-Lite
AR	**coNP**	**coNP**
Brave	**NP**	**NL**

Example: handling inconsistency

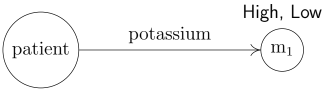
TBox $\mathcal{T}$:

$$\mathrm{High} \sqcap \mathrm{Low} \sqsubseteq \bot$$

$$\exists \mathrm{potassium}.\mathrm{Low} \sqcap \mathrm{CardiacDisease} \sqsubseteq \mathrm{CardiacArrest}$$
$$\exists \mathrm{potassium}.\mathrm{High} \sqcap \mathrm{CardiacDisease} \sqsubseteq \mathrm{CardiacArrest}$$

$$\exists \mathrm{potassium}.\mathrm{High} \sqcap \exists \mathrm{calcium}.\mathrm{Low} \sqsubseteq \mathrm{CardiacArrest}$$

ABox $\mathcal{A}$:



- observation: $\mathrm{CardiacArrest}(\mathrm{patient})$
- repairs: remove either $\mathrm{High}(\mathrm{m}_1)$ or $\mathrm{Low}(\mathrm{m}_1)$
- Brave hypotheses: $\{\mathrm{CardiacDisease}(\mathrm{patient})\}$ and $\{\mathrm{Low}(\mathrm{m}_2), \mathrm{calcium}(\mathrm{patient}, \mathrm{m}_2)\}$
- AR hypothesis: $\{\mathrm{CardiacDisease}(\mathrm{patient})\}$

ABox abduction for inconsistent KBs

Definition

Inconsistent KB $\mathcal{K}$, concept assertion α (observation) s.t. $\mathcal{K} \not\models_{\mathrm{AR\ (Brave)}} \alpha$.

- $\langle \mathcal{K}, \alpha \rangle$ called AR-abduction problem (Brave-abduction problem)
- AR-hypothesis (Brave-hypothesis) for $\langle \mathcal{K}, \alpha \rangle$: ABox $\mathcal{H}$ s.t. $\mathcal{K} \cup \mathcal{H} \models_{\mathrm{AR\ (Brave)}} \alpha$

properties of desirable hypotheses

- non-trivial: $\alpha \notin \mathcal{H}$
- signature-restricted: only names from given signature
- conflict-confining: same set of conflicts $(\mathrm{Conf}(\mathcal{K} \cup \mathcal{H}) = \mathrm{Conf}(\mathcal{K}))$

ABox abduction for inconsistent KBs

Definition

Inconsistent KB $\mathcal{K}$, concept assertion α (observation) s.t. $\mathcal{K} \not\models_{\mathrm{AR\ (Brave)}} \alpha$.

- $\langle \mathcal{K}, \alpha \rangle$ called AR-abduction problem (Brave-abduction problem)
- AR-hypothesis (Brave-hypothesis) for $\langle \mathcal{K}, \alpha \rangle$: ABox $\mathcal{H}$ s.t. $\mathcal{K} \cup \mathcal{H} \models_{\mathrm{AR\ (Brave)}} \alpha$

properties of desirable hypotheses (change as little as possible)

- $\subseteq$-minimal: subset-minimal (w.r.t. $\mathcal{H}$ itself)
- $\leq$-minimal: cardinality-minimal (w.r.t. $\mathcal{H}$ itself)
- $\subseteq_c$-minimal: subset-minimal w.r.t. set of conflicts
- $\leq_c$-minimal: cardinality-minimal w.r.t. set of conflicts

Decision problems for ABox abduction

- existence problem: Given an AR-abduction $(\mathrm{Brave}$-abduction$)$ problem $\langle\mathcal{K},\alpha\rangle$, is there an AR-hypothesis $(\mathrm{Brave}$-hypothesis$)$ for $\langle\mathcal{K},\alpha\rangle$?

- verification problem: Given an AR-abduction $(\mathrm{Brave}$-abduction$)$ problem $\langle\mathcal{K},\alpha\rangle$ and ABox $\mathcal{H}$, is $\mathcal{H}$ an AR-hypothesis $(\mathrm{Brave}$-hypothesis$)$ for $\langle\mathcal{K},\alpha\rangle$?

Non-convexity

PADERBORN UNIVERSITY

- set of AR-hypotheses for an AR-abduction problem can be non-convex: $\mathcal{B}_1 \subsetneq \mathcal{B}_2 \subsetneq \mathcal{B}_3$, where $\mathcal{B}_1$ and $\mathcal{B}_3$ are AR-hypotheses, but $\mathcal{B}_2$ is not
- $\subseteq$-minimality cannot be checked locally by looking at direct subsets

$$\mathcal{T} = \{C_1 \sqcap C_2 \sqsubseteq \bot, \quad B_1 \sqsubseteq A, \quad B_1 \sqcap B_2 \sqsubseteq \bot, \quad B_3 \sqsubseteq A\}$$

$$\mathcal{A} = \{C_1(m), C_2(m)\}$$

$$\alpha = A(m)$$

$$\mathcal{B}_1 = \{B_1(m)\} \qquad \mathcal{B}_2 = \mathcal{B}_1 \cup \{B_2(m)\} \qquad \mathcal{B}_3 = \mathcal{B}_2 \cup \{B_3(m)\}$$

Non-convexity and $\subseteq$-minimality for $\mathcal{EL}_\bot$

$$\mathcal{T}_0 \coloneqq \{\, B_1 \sqcap B_2 \sqsubseteq \bot, \quad B_1' \sqcap B_2 \sqsubseteq \bot \,\}$$

$$\mathcal{T}_1 \coloneqq \left\{ B_1 \sqcap \bigsqcap_{1 \leq i \leq n} (T_{x,i} \sqcap F_{x,i}) \sqsubseteq A \right\} \cup$$

$$\{\, B_1 \sqcap C_j \sqsubseteq A \mid 1 \leq j \leq k \,\} \cup \{\, B_1' \sqcap T_{y,i} \sqcap F_{y,i} \sqsubseteq \bot \mid 1 \leq i \leq m \,\} \cup$$

$$\{\, B_1' \sqcap T_{x,i} \sqcap C_j \sqsubseteq \bot \mid \neg x_i \in c_j \,\} \cup \{\, B_1' \sqcap F_{x,i} \sqcap C_j \sqsubseteq \bot \mid x_i \in c_j \,\} \cup$$

$$\{\, B_1' \sqcap T_{y,i} \sqcap C_j \sqsubseteq \bot \mid \neg y_i \in c_j \,\} \cup \{\, B_1' \sqcap F_{y,i} \sqcap C_j \sqsubseteq \bot \mid y_i \in c_j \,\}$$

$$\mathcal{T}_2 \coloneqq \left\{ B_2 \sqcap \bigsqcap_{1 \leq i \leq n} \mathrm{HAVE}_i \sqsubseteq A \right\} \cup \{\, T_{x,i} \sqsubseteq \mathrm{HAVE}_i,\ F_{x,i} \sqsubseteq \mathrm{HAVE}_i \mid 1 \leq i \leq n \,\}$$

$$\mathcal{A} \coloneqq \{B_1(a), B_1'(a), B_2(a)\} \cup \{\, T_{y,i}(a), F_{y,i}(a) \mid 1 \leq i \leq m \,\} \cup \{\, C_j(a) \mid 1 \leq j \leq k \,\}$$

Non-convexity and $\subseteq$-minimality for DL-Lite

- DL-Lite has no conjunction
- minimal $\mathcal{T}$-supports of concept assertions α are of size 1:
- $\mathcal{T} \cup \mathcal{A} \models \alpha \quad \Longrightarrow \quad \mathcal{T} \cup \{\beta\} \models \alpha$ for some $\beta \in \mathcal{A}$

Lemma

For AR-abduction problem $\langle \mathcal{K}, \alpha \rangle$ and ABox $\mathcal{B}$, the following are equivalent:

1. *there is an AR-hypothesis $\mathcal{H} \subseteq \mathcal{B}$ for $\langle \mathcal{K}, \alpha \rangle$*
2. *$\{\mathcal{R}_1, \ldots, \mathcal{R}_\ell\}$ repairs of $\mathcal{K}$ $\rightsquigarrow$ there are $\mathcal{T}$-supports $\{\beta_i\} \subseteq \mathcal{A} \cup \mathcal{B}$ of α s.t. $\mathcal{R}_i \cup \{\beta_i\}$ is $\mathcal{T}$-consistent for each i.*

Proof idea.

"1. $\Rightarrow$ 2.": each repair $\mathcal{R}$ of $\mathcal{T} \cup \mathcal{A} \cup \mathcal{H}$ entails α, giving us one of the β_i

"2. $\Rightarrow$ 1.": use set of all β_i $\rightsquigarrow$ each repair of the new KB contains some β_i $\square$

Non-convexity and $\subseteq$-minimality for DL-Lite

Lemma

For AR-abduction problem $\langle\mathcal{K},\alpha\rangle$ and ABox $\mathcal{B}$, the following are equivalent:

1. *there is an AR-hypothesis $\mathcal{H} \subseteq \mathcal{B}$ for $\langle\mathcal{K},\alpha\rangle$*
2. *$\{\mathcal{R}_1,\ldots,\mathcal{R}_\ell\}$ repairs of $\mathcal{K}$ $\rightsquigarrow$ there are $\mathcal{T}$-supports $\{\beta_i\} \subseteq \mathcal{A} \cup \mathcal{B}$ of α s.t. $\mathcal{R}_i \cup \{\beta_i\}$ is $\mathcal{T}$-consistent for each i.*

- given $\mathcal{T} \cup \mathcal{A}$ and $\mathcal{H}$, want to check for an AR-hypothesis $\mathcal{H}' \subsetneq \mathcal{H}$
- let $\mathcal{H} = \{\gamma_1,\ldots,\gamma_n\}$ and $\mathcal{B}_i \coloneqq \mathcal{H} \setminus \{\gamma_i\}$
- for each $\mathcal{B}_i$, check for an AR-hypothesis $\mathcal{H}' \subseteq \mathcal{B}_i$
- use the lemma! (for each i)
- $\forall$ repairs $\mathcal{R}$, iterate over all $\beta \in \mathcal{A} \cup \mathcal{B}_i$
- check: $\mathcal{T} \cup \{\beta\} \models \alpha$ and $\mathcal{R} \cup \{\beta\}$ is $\mathcal{T}$-consistent

Results for $\mathcal{EL}_\perp$

existence problem	general	signature	conflict-confining	non-trivial
Brave	P	NP	Σ_2^{P}	NP
AR	coNP	Σ_2^{P}	coNP	Σ_2^{P}

verification problem	$\leq$-min.	$\subseteq$-min.	$\subseteq_c$-min.	$\leq_c$-min.	conflict-confining
Brave	NP	DP	Π_2^{P}	–	DP
AR	coNP	Π_2^{P}	coNP	coNP	coNP

Results for DL-Lite

PADERBORN UNIVERSITY

existence problem	general	signature	conflict-confining	non-trivial
Brave	NL	NL	NL	NL
AR	NL	coNP	NL	coNP

verification problem	$\leq$-min.	$\subseteq$-min.	$\subseteq_c$-min.	$\leq_c$-min.	conflict-confining
Brave	NL	NL	NL	NL	NL
AR	coNP	DP	NL	NL	NL

Summary

- systematic study of ABox abduction under repair semantics
  - lightweight DLs $\mathcal{EL}_\bot$ and $\mathrm{DL\text{-}Lite}$
  - introduced interesting new properties of explanations in this setting
- obtained first results on combined complexity
- found interesting effects for abduction in our setting
  - non-convexity: second level of **PH** for $\mathcal{EL}_\bot$, but not $\mathrm{DL\text{-}Lite}$
  - existence of conflict-confining Brave hypotheses

Outlook and future work

- complexity for combinations of properties in $\mathcal{EL}_\bot$ at DL '26
- many variations, e.g. other semantics, new properties
- data complexity
- more expressive observations, such as (conjunctive) queries