

INCONSISTENCY-TOLERANT QUERYING OF DESCRIPTION LOGIC KNOWLEDGE BASES

Meghyn Bienvenu (*CNRS, University of Montpellier, Inria*)

ONTOLOGY-MEDIATED QUERY ANSWERING (OMQA)



**incomplete
database**
(ground facts)



ontology
(logical theory)



user query

ONTOLOGY-MEDIATED QUERY ANSWERING (OMQA)



patient data

“Melanie has listeriosis”
“Paul has Lyme disease”



medical knowledge

“Listeriosis & Lyme disease
are bacterial infections”



user query

“Find all patients with
bacterial infections”

ONTOLOGY-MEDIATED QUERY ANSWERING (OMQA)



patient data

“Melanie has listeriosis”
“Paul has Lyme disease”



medical knowledge

“Listeriosis & Lyme disease
are bacterial infections”



user query

“Find all patients with
bacterial infections”

expected answers: Melanie, Paul

ONTOLOGY-MEDIATED QUERY ANSWERING (OMQA)



employee data

“Marie is a professor”
“Mark teaches CS200”



org. knowledge

“Professors are teaching staff”
“Someone who teaches is
part of the teaching staff”



user query

“Find all teaching staff”

ONTOLOGY-MEDIATED QUERY ANSWERING (OMQA)



employee data

“Marie is a professor”
“Mark teaches CS200”



org. knowledge

“Professors are teaching staff”
“Someone who teaches is part of the teaching staff”



user query

“Find all teaching staff”

expected answers: Marie, Mark

WHAT ARE ONTOLOGIES GOOD FOR?

To **standardize the terminology** of an application domain

- by adopting a common vocabulary, **easy to share information**
- **meaning of terms is constrained**, so less misunderstandings

WHAT ARE ONTOLOGIES GOOD FOR?

To **standardize the terminology** of an application domain

- by adopting a common vocabulary, **easy to share information**
- **meaning of terms is constrained**, so less misunderstandings

To present an **intuitive and unified view of data sources**

- ontology can be used to **enrich the data vocabulary**, making it **easier for users to formulate their queries**
- especially useful when **integrating multiple data sources**

WHAT ARE ONTOLOGIES GOOD FOR?

To **standardize the terminology** of an application domain

- by adopting a common vocabulary, **easy to share information**
- **meaning of terms is constrained**, so less misunderstandings

To present an **intuitive and unified view of data sources**

- ontology can be used to **enrich the data vocabulary**, making it **easier for users to formulate their queries**
- especially useful when **integrating multiple data sources**

To support **automated reasoning**

- **uncover implicit connections** between terms, **errors in modelling**
- **exploit knowledge in the ontology during query answering**, to get back a **more complete set of answers** to queries

Description logics (DLs):

- popular means for specifying ontologies
- basis of the web ontology language OWL (W3C standard)

Formally: **decidable fragments of first-order logic**

- inherit well-defined semantics
- succinct, variable-free syntax

Description logics (DLs):

- popular means for specifying ontologies
- basis of the web ontology language OWL (W3C standard)

Formally: **decidable fragments of first-order logic**

- inherit well-defined semantics
- succinct, variable-free syntax

Computational properties well understood (decidability, complexity)

Many **implemented reasoners and tools** available for use

In realistic settings, can expect some **errors in the data**

- ABox likely to be **inconsistent** with the TBox (ontology)

Standard semantics: everything is implied - **not informative!**

In realistic settings, can expect some **errors in the data**

- ABox likely to be **inconsistent** with the TBox (ontology)

Standard semantics: everything is implied - **not informative!**

Two approaches to inconsistency handling:

- **resolve the inconsistencies**
 - preferable, but not always applicable!
- live with the inconsistencies - **adopt alternative semantics**
 - **meaningful answers** to queries despite inconsistencies

Note: focus on case where errors in ABox (**assume TBox reliable**)

INTRODUCTION TO DLS & OMQA

Building blocks:

- **concept names** (unary predicates, classes)
- **role names** (binary predicates, properties)

Prof Fac Course

Teaches HeadOf

Constructors to build complex descriptions

$\sqcup, \sqcap, \neg, \forall, \exists, \dots$

$\text{Fac} \sqcap \neg \text{Prof}$

$\exists \text{Teaches. GradCourse}$

Teaches^-

Building blocks:

- **concept names** (unary predicates, classes)
- **role names** (binary predicates, properties)

Prof Fac Course

Teaches HeadOf

Constructors to build complex descriptions

$\sqcup, \sqcap, \neg, \forall, \exists, \dots$

$\text{Fac} \sqcap \neg\text{Prof}$

$\exists\text{Teaches}.\text{GradCourse}$

Teaches^-

TBox (ontology) = set of axioms

- **concept inclusions**

$\text{Prof} \sqsubseteq \text{Fac}$

$\text{Prof} \sqsubseteq \neg\text{Fellow}$

$\exists\text{Teaches}.\text{GradCourse} \sqsubseteq \text{Prof}$

- **role inclusions**

$\text{TaughtBy} \sqsubseteq \text{Teaches}^-$

$\text{HeadOf} \sqsubseteq \text{MemberOf}$

Note: allowed constructors and axioms **depends on chosen DL**

Mainly focus on DLs of the DL-Lite family

- specifically designed for OMQA
- simple DLs with useful modelling constructs
- basis for OWL 2 QL profile

Mainly focus on DLs of the DL-Lite family

- specifically designed for OMQA
- simple DLs with useful modelling constructs
- basis for OWL 2 QL profile

DL-Lite_R dialect:

- concept inclusions $B_1 \sqsubseteq (\neg)B_2$ B_1, B_2 either $A \in N_C$ or $\exists R (R \in N_R^\pm)$
- role inclusions $R_1 \sqsubseteq (\neg)R_2$ $R_1, R_2 \in N_R^\pm$

We will also briefly consider other DLs

‘Lightweight’ description logic \mathcal{EL}_{\perp} :

- concept constructors: \top, \perp, \sqcap , and $\exists r.C$
- only concept inclusions $C \sqsubseteq D$ in TBox

‘Expressive’ description logic \mathcal{ALC} :

- concept constructors: $\top, \perp, \neg, \sqcap, \sqcup, \exists r.C$, and $\forall r.C$
- only concept inclusions $C \sqsubseteq D$ in TBox

Interpretation \mathcal{I} (“possible world”)

- **domain of objects** $\Delta^{\mathcal{I}}$ (possibly infinite set)
- **interpretation function** $\cdot^{\mathcal{I}}$ that maps
 - **concept name** $A \rightsquigarrow$ set of objects $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
 - **role name** $r \rightsquigarrow$ set of pairs of objects $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
 - **individual name** $a \rightsquigarrow$ object $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
- **extend** $\cdot^{\mathcal{I}}$ to **complex concepts and roles** in natural way

Interpretation \mathcal{I} (“possible world”)

- **domain of objects** $\Delta^{\mathcal{I}}$ (possibly infinite set)
- **interpretation function** $\cdot^{\mathcal{I}}$ that maps
 - **concept name** $A \rightsquigarrow$ set of objects $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
 - **role name** $r \rightsquigarrow$ set of pairs of objects $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
 - **individual name** $a \rightsquigarrow$ object $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
- **extend** $\cdot^{\mathcal{I}}$ to **complex concepts and roles** in natural way

Interpretation \mathcal{I} is a **model** of KB $\langle \mathcal{T}, \mathcal{A} \rangle$ if:

- $G^{\mathcal{I}} \subseteq H^{\mathcal{I}}$ for every (concept or role) inclusion $G \sqsubseteq H \in \mathcal{T}$
- $a^{\mathcal{I}} \in A^{\mathcal{I}}$ for every $A(a) \in \mathcal{A}$ and $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$ for every $r(a, b) \in \mathcal{A}$

Satisfiable KB = has **at least one model**

Interpretation \mathcal{I} (“possible world”)

- **domain of objects** $\Delta^{\mathcal{I}}$ (possibly infinite set)
- **interpretation function** $\cdot^{\mathcal{I}}$ that maps
 - **concept name** $A \rightsquigarrow$ set of objects $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
 - **role name** $r \rightsquigarrow$ set of pairs of objects $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
 - **individual name** $a \rightsquigarrow$ object $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
- **extend** $\cdot^{\mathcal{I}}$ to **complex concepts and roles** in natural way

Interpretation \mathcal{I} is a **model** of KB $\langle \mathcal{T}, \mathcal{A} \rangle$ if:

- $G^{\mathcal{I}} \subseteq H^{\mathcal{I}}$ for every (concept or role) inclusion $G \sqsubseteq H \in \mathcal{T}$
- $a^{\mathcal{I}} \in A^{\mathcal{I}}$ for every $A(a) \in \mathcal{A}$ and $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$ for every $r(a, b) \in \mathcal{A}$

Satisfiable KB = has **at least one model**

ABox \mathcal{A} is \mathcal{T} -consistent = KB $\langle \mathcal{T}, \mathcal{A} \rangle$ is satisfiable

EXAMPLE: UNSATISFIABLE KB

Example TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Lect}$	$\text{Fac} \sqsubseteq \neg \text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists \text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg \text{Fellow}$	

Example ABox $\mathcal{A}_{\text{univ}}$:

Prof(anna), Lect(anna), Fellow(anna), Prof(kim), Lect(kim),
Fellow(julie), Teaches(csc343, julie), Fellow(alex), Teaches(alex, csc486)

Claim: $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ is unsatisfiable

\sqsubseteq -minimal $\mathcal{T}_{\text{univ}}$ -inconsistent subsets of $\mathcal{A}_{\text{univ}}$:

EXAMPLE: UNSATISFIABLE KB

Example TBox $\mathcal{T}_{\text{univ}}$:

Prof \sqsubseteq Fac Prof $\sqsubseteq \exists\text{Teaches}$ Prof $\sqsubseteq \neg\text{Lect}$ Fac $\sqsubseteq \neg\text{Course}$
Lect \sqsubseteq Fac Lect $\sqsubseteq \exists\text{Teaches}$ Prof $\sqsubseteq \neg\text{Fellow}$
Fellow \sqsubseteq Fac $\exists\text{Teaches}^- \sqsubseteq \text{Course}$ Lect $\sqsubseteq \neg\text{Fellow}$

Example ABox $\mathcal{A}_{\text{univ}}$:

Prof(anna), Lect(anna), Fellow(anna), Prof(kim), Lect(kim),
Fellow(julie), Teaches(csc343, julie), Fellow(alex), Teaches(alex, csc486)

Claim: $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ is unsatisfiable

\sqsubseteq -minimal $\mathcal{T}_{\text{univ}}$ -inconsistent subsets of $\mathcal{A}_{\text{univ}}$:

{Prof(anna), Lect(anna)} {Prof(anna), Fellow(anna)}
{Lect(anna), Fellow(anna)} {Prof(kim), Lect(kim)}
{Fellow(julie), Teaches(csc343, julie)}

Instance queries (IQs): find instances of a given concept or role

Fac(x)

Teaches(x,y)

Instance queries (IQs): find instances of a given concept or role

Fac(x)

Teaches(x,y)

Conjunctive queries (CQs) ~ SPJ queries in SQL, BGPs in SPARQL
conjunctions of atoms, some variables can be existentially quantified

$\exists y. \text{Fac}(x) \wedge \text{Teaches}(x,y)$

(find all faculty members that teach something)

Instance queries (IQs): find instances of a given concept or role

Fac(x)

Teaches(x,y)

Conjunctive queries (CQs) ~ SPJ queries in SQL, BGPs in SPARQL
 conjunctions of atoms, some variables can be existentially quantified

$\exists y. \text{Fac}(x) \wedge \text{Teaches}(x,y)$

(find all faculty members that teach something)

Unions of conjunctive queries (UCQs): disjunction of CQs

Answers to $q(\vec{x})$ in interpretation \mathcal{I} :

tuples $\vec{o} \in \Delta^{\mathcal{I}}$ such that $q(\vec{x} \mapsto \vec{o})$ holds in \mathcal{I}

Answers to $q(\vec{x})$ in interpretation \mathcal{I} :

tuples $\vec{o} \in \Delta^{\mathcal{I}}$ such that $q(\vec{x} \mapsto \vec{o})$ holds in \mathcal{I}

Problem: **each KB** gives rise to **multiple interpretations** (its models)

Answers to $q(\vec{x})$ in interpretation \mathcal{I} :

tuples $\vec{o} \in \Delta^{\mathcal{I}}$ such that $q(\vec{x} \mapsto \vec{o})$ holds in \mathcal{I}

Problem: each KB gives rise to **multiple interpretations** (its models)

Solution: adopt **certain answer semantics**

- require tuple to be an answer w.r.t. **all models of KB**

Answers to $q(\vec{x})$ in interpretation \mathcal{I} :

tuples $\vec{o} \in \Delta^{\mathcal{I}}$ such that $q(\vec{x} \mapsto \vec{o})$ holds in \mathcal{I}

Problem: each KB gives rise to **multiple interpretations** (its models)

Solution: adopt **certain answer semantics**

- require tuple to be an answer w.r.t. **all models of KB**

Formally: Call a tuple (a_1, \dots, a_n) of individuals from \mathcal{A} a **certain answer to n -ary query q over DL KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$** iff

$(a_1^{\mathcal{I}}, \dots, a_n^{\mathcal{I}}) \in \text{ans}(q, \mathcal{I})$ for every model \mathcal{I} of \mathcal{K}

Notation: $\mathcal{K} \models q(a_1, \dots, a_n)$

Answers to $q(\vec{x})$ in interpretation \mathcal{I} :

tuples $\vec{o} \in \Delta^{\mathcal{I}}$ such that $q(\vec{x} \mapsto \vec{o})$ holds in \mathcal{I}

Problem: each KB gives rise to multiple interpretations (its models)

Solution: adopt certain answer semantics

- require tuple to be an answer w.r.t. all models of KB

Formally: Call a tuple (a_1, \dots, a_n) of individuals from \mathcal{A} a certain answer to n -ary query q over DL KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ iff

$$(a_1^{\mathcal{I}}, \dots, a_n^{\mathcal{I}}) \in \text{ans}(q, \mathcal{I}) \text{ for every model } \mathcal{I} \text{ of } \mathcal{K}$$

Notation: $\mathcal{K} \models q(a_1, \dots, a_n)$

Ontology-mediated query answering (OMQA)

= computing certain answers to queries

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

Prof(anna), Lect(kim), Fellow(julie), Fellow(alex), Teaches(alex, csc486)

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_1(x) = \text{Fac}(x)$

Certain answers to q_1 :

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_1(x) = \text{Fac}(x)$

Certain answers to q_1 : $\text{anna}, \text{kim}, \text{julie}, \text{alex}$

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_2(x) = \exists y \text{Teaches}(x, y)$

Certain answers to q_2 :

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_2(x) = \exists y \text{Teaches}(x, y)$

Certain answers to q_2 : $\text{anna}, \text{kim}, \text{alex}$

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_3(x) = \exists y \text{Fac}(x) \wedge \text{Teaches}(x, y)$

Certain answers to q_3 :

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_3(x) = \exists y \text{Fac}(x) \wedge \text{Teaches}(x, y)$

Certain answers to q_3 : $\text{anna}, \text{kim}, \text{alex}$

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_4(x, y) = \text{Fac}(x) \wedge \text{Teaches}(x, y)$

Certain answers to q_4 :

EXAMPLE: OMQA WITH DL-LITE ONTOLOGIES

Same TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Lect}$	$\text{Fac} \sqsubseteq \neg\text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists\text{Teaches}$	$\text{Prof} \sqsubseteq \neg\text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists\text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg\text{Fellow}$	

Consistent subset of $\mathcal{A}_{\text{univ}}$:

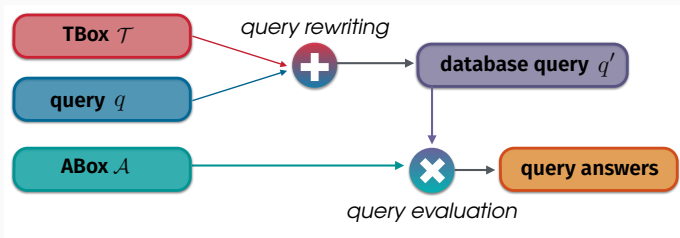
$\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Query: $q_4(x, y) = \text{Fac}(x) \wedge \text{Teaches}(x, y)$

Certain answers to q_4 : $(\text{alex}, \text{csc486})$

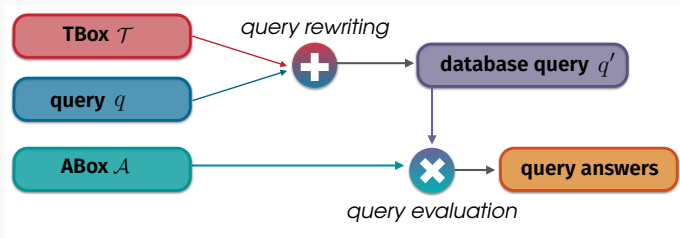
KEY TECHNIQUE FOR OMQA: QUERY REWRITING

Query rewriting: **Reduces** problem of **finding certain answers** to standard DB query evaluation (↪ exploit existing DB systems)



KEY TECHNIQUE FOR OMQA: QUERY REWRITING

Query rewriting: **Reduces** problem of **finding certain answers** to standard DB query evaluation (↪ exploit existing DB systems)

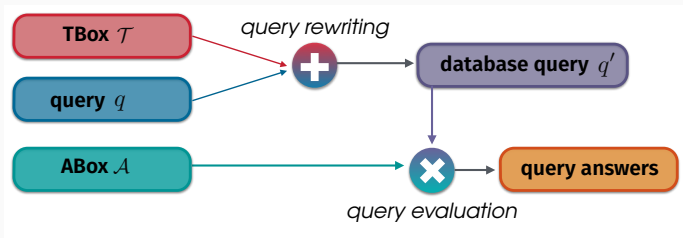


Call $q'(\vec{x})$ a **rewriting** of $q(\vec{x})$ and \mathcal{T} iff for every ABox \mathcal{A} and tuple \vec{a}

$$\mathcal{T}, \mathcal{A} \models q(\vec{a}) \quad \Leftrightarrow \quad \vec{a} \in \text{ans}(q'(\vec{x}), \mathcal{I}_{\mathcal{A}}) \quad (\mathcal{I}_{\mathcal{A}} = \text{treat } \mathcal{A} \text{ as DB})$$

KEY TECHNIQUE FOR OMQA: QUERY REWRITING

Query rewriting: **Reduces** problem of **finding certain answers** to standard DB query evaluation (\rightsquigarrow exploit existing DB systems)



Call $q'(\vec{x})$ a **rewriting** of $q(\vec{x})$ and \mathcal{T} iff for every ABox \mathcal{A} and tuple \vec{a}

$$\mathcal{T}, \mathcal{A} \models q(\vec{a}) \Leftrightarrow \vec{a} \in \text{ans}(q'(\vec{x}), \mathcal{I}_{\mathcal{A}}) \quad (\mathcal{I}_{\mathcal{A}} = \text{treat } \mathcal{A} \text{ as DB})$$

First-order (FO) rewritings: q' is an FO (\sim SQL) query

UCQ-rewritings: q' is a UCQ

EXAMPLE: QUERY REWRITING

Rewriting of $q_1(x) = \text{Fac}(x)$ w.r.t. $\mathcal{T}_{\text{univ}}$:

$$q'_1(x) = \text{Fac}(x) \vee \text{Prof}(x) \vee \text{Lect}(x) \vee \text{Fellow}(x)$$

Rewriting of $q_3(x) = \exists y \text{Fac}(x) \wedge \text{Teaches}(x, y)$ w.r.t. $\mathcal{T}_{\text{univ}}$:

$$q'_3(x) = (\exists y. \text{Fac}(x) \wedge \text{Teaches}(x, y)) \vee \text{Prof}(x) \vee \text{Lect}(x) \vee (\exists y. \text{Fellow}(x) \wedge \text{Teaches}(x, y))$$

Rewriting of $q_4(x, y) = \text{Fac}(x) \wedge \text{Teaches}(x, y)$ w.r.t. $\mathcal{T}_{\text{univ}}$:

$$q'_4(x, y) = (\text{Fac}(x) \wedge \text{Teaches}(x, y)) \vee (\text{Prof}(x) \wedge \text{Teaches}(x, y)) \vee (\text{Lect}(x) \wedge \text{Teaches}(x, y)) \vee (\text{Fellow}(x) \wedge \text{Teaches}(x, y))$$

INTRODUCTION TO INCONSISTENCY- TOLERANT SEMANTICS

In realistic settings, can expect some **errors in the data**

- ABox likely to be **inconsistent** with the TBox (ontology)

Standard semantics: everything is implied - **not informative!**

In realistic settings, can expect some **errors in the data**

- ABox likely to be **inconsistent** with the TBox (ontology)

Standard semantics: everything is implied - **not informative!**

Two approaches to inconsistency handling:

- **resolve the inconsistencies**
 - preferable, but not always applicable!
- live with the inconsistencies - **adopt alternative semantics**
 - **meaningful answers** to queries despite inconsistencies

Note: focus on case where errors in ABox (**assume TBox reliable**)

EXAMPLE: REASONABLE INFERENCES

TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Lect}$	$\text{Fac} \sqsubseteq \neg \text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists \text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg \text{Fellow}$	

Consider following ABoxes:

\mathcal{A}_1	=	{Prof(anna), Lect(anna), Fellow(alex)}
\mathcal{A}_2	=	{Prof(anna), Fellow(alex), Lect(alex)}

Which assertions would be reasonable to infer from these two KBs?

Prof(anna)	Lect(anna)	Fac(anna)
Fellow(alex)	Lect(alex)	Fac(alex)

EXAMPLE: REASONABLE ANSWERS

TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Lect}$	$\text{Fac} \sqsubseteq \neg \text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists \text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg \text{Fellow}$	

ABox $\mathcal{A}_{\text{univ}}$:

Prof(anna), Lect(anna), Fellow(anna), Prof(kim), Lect(kim),
Fellow(julie), Teaches(csc343, julie), Fellow(alex), Teaches(alex, csc486)

Question: what are **reasonable answers** for our example queries?

$$q_1(x) = \text{Fac}(x)$$

$$q_2(x) = \exists y \text{Teaches}(x, y)$$

$$q_3(x) = \exists y \text{Fac}(x) \wedge \text{Teaches}(x, y)$$

$$q_4(x, y) = \text{Fac}(x) \wedge \text{Teaches}(x, y)$$

In general: **no single best way** to define answers for inconsistent KBs
⇒ consider **many different inconsistency-tolerant semantics**

Formally: a **semantics \mathcal{S}** associates a set of **query answers** to **every KB and query**

- if \mathcal{K} is **satisfiable**, should return **certain answers**
- for **unsatisfiable** \mathcal{K} , can give **different answers** than classical semantics

Write $\mathcal{K} \models_{\mathcal{S}} q(\vec{a})$ if \vec{a} **answer to q w.r.t. \mathcal{K} under semantics \mathcal{S}**

Consider different ways of **comparing semantics**

Call $C \subseteq \mathcal{A}$ is a **(consistent) \mathcal{T} -support** of $q(\vec{a})$ if:

(i) C is **\mathcal{T} -consistent** (ii) $\langle \mathcal{T}, C \rangle \models q(\vec{a})$

Semantics S satisfies the **CONSISTENT SUPPORT property** if whenever $\mathcal{K} \models_s q(\vec{a})$, there **exists a \mathcal{T} -support $C \subseteq \mathcal{A}$ of $q(\vec{a})$**

- important for **explaining / justifying query results to users**

Call $C \subseteq \mathcal{A}$ is a **(consistent) \mathcal{T} -support** of $q(\vec{a})$ if:

- (i) C is **\mathcal{T} -consistent** (ii) $\langle \mathcal{T}, C \rangle \models q(\vec{a})$

Semantics S satisfies the **CONSISTENT SUPPORT property** if whenever $\mathcal{K} \models_S q(\vec{a})$, there **exists a \mathcal{T} -support $C \subseteq \mathcal{A}$ of $q(\vec{a})$**

- important for **explaining / justifying query results to users**

Semantics S satisfies the **CONSISTENT RESULTS property** if for every KB \mathcal{K} , there exists a model \mathcal{I} of \mathcal{T} such that $\mathcal{K} \models_S q(\vec{a})$ implies $\mathcal{I} \models q(\vec{a})$.

- set of **query results** is **jointly consistent with TBox**
- **safe to combine query results**

Call $C \subseteq \mathcal{A}$ is a **(consistent) \mathcal{T} -support** of $q(\vec{a})$ if:

- (i) C is **\mathcal{T} -consistent** (ii) $\langle \mathcal{T}, C \rangle \models q(\vec{a})$

Semantics S satisfies the **CONSISTENT SUPPORT property** if whenever $\mathcal{K} \models_S q(\vec{a})$, there **exists a \mathcal{T} -support $C \subseteq \mathcal{A}$ of $q(\vec{a})$**

- important for **explaining / justifying query results to users**

Semantics S satisfies the **CONSISTENT RESULTS property** if for every KB \mathcal{K} , there exists a model \mathcal{I} of \mathcal{T} such that $\mathcal{K} \models_S q(\vec{a})$ implies $\mathcal{I} \models q(\vec{a})$.

- set of **query results** is **jointly consistent with TBox**
- **safe to combine query results**

Note: **neither property implies the other**

Given two semantics S and S' , we say that:

- S' is an **under-approximation** (or: **sound approximation**) of S just in the case that

$$\mathcal{K} \models_{S'} q(\vec{a}) \Rightarrow \mathcal{K} \models_S q(\vec{a})$$

- S' is an **over-approximation** (or: **complete approximation**) of S just in the case that

$$\mathcal{K} \models_S q(\vec{a}) \Rightarrow \mathcal{K} \models_{S'} q(\vec{a})$$

Consistency properties are **preserved by under-approximations**:

S' is an **under-approximation** of S & S satisfies $P \Rightarrow S'$ also satisfies P

here $P \in \{\text{CONSISTENT SUPPORT, CONSISTENT RESULTS}\}$

Many semantics are based upon the notion of repair

Repair of an ABox \mathcal{A} w.r.t. a TBox \mathcal{T}

= inclusion-maximal subset of \mathcal{A} that is \mathcal{T} -consistent

Intuition: different **ways of achieving consistency** while **retaining as much of the original data as possible**

Denote by $Rep(\mathcal{A}, \mathcal{T})$ the **set of repairs of \mathcal{A} w.r.t. \mathcal{T}**

- abbreviate to $Rep(\mathcal{K})$ when $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$

Every KB has **at least one repair**

- inconsistent KB \Rightarrow **typically multiple repairs**

EXAMPLE: REPAIRS

Reconsider the TBox $\mathcal{T}_{\text{univ}}$:

$\text{Prof} \sqsubseteq \text{Fac}$	$\text{Prof} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Lect}$	$\text{Fac} \sqsubseteq \neg \text{Course}$
$\text{Lect} \sqsubseteq \text{Fac}$	$\text{Lect} \sqsubseteq \exists \text{Teaches}$	$\text{Prof} \sqsubseteq \neg \text{Fellow}$	
$\text{Fellow} \sqsubseteq \text{Fac}$	$\exists \text{Teaches}^- \sqsubseteq \text{Course}$	$\text{Lect} \sqsubseteq \neg \text{Fellow}$	

and ABox $\mathcal{A}_{\text{univ}}$:

$\text{Prof}(\text{anna}), \text{Lect}(\text{anna}), \text{Fellow}(\text{anna}), \text{Prof}(\text{kim}), \text{Lect}(\text{kim}),$
 $\text{Fellow}(\text{julie}), \text{Teaches}(\text{csc343}, \text{julie}), \text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})$

Recall the minimal $\mathcal{T}_{\text{univ}}$ -inconsistent subsets:

$\{\text{Prof}(\text{anna}), \text{Lect}(\text{anna})\}$	$\{\text{Prof}(\text{anna}), \text{Fellow}(\text{anna})\}$
$\{\text{Lect}(\text{anna}), \text{Fellow}(\text{anna})\}$	$\{\text{Prof}(\text{kim}), \text{Lect}(\text{kim})\}$
$\{\text{Fellow}(\text{julie}), \text{Teaches}(\text{csc343}, \text{julie})\}$	

Question: **How many repairs of $\mathcal{A}_{\text{univ}}$ w.r.t. $\mathcal{T}_{\text{univ}}$?**

EXAMPLE: REPAIRS (CONT.)

Twelve repairs of $\mathcal{A}_{\text{univ}}$ w.r.t. $\mathcal{T}_{\text{univ}}$:

$$\begin{aligned}\mathcal{R}_1 &= \{\text{Prof}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_2 &= \{\text{Lect}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_3 &= \{\text{Fellow}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_4 &= \{\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_5 &= \{\text{Lect}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_6 &= \{\text{Fellow}(\text{anna}), \text{Lect}(\text{kim}), \text{Fellow}(\text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_7 &= \{\text{Prof}(\text{anna}), \text{Prof}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_8 &= \{\text{Lect}(\text{anna}), \text{Lect}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_9 &= \{\text{Fellow}(\text{anna}), \text{Prof}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_{10} &= \{\text{Prof}(\text{anna}), \text{Lect}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_{11} &= \{\text{Lect}(\text{anna}), \text{Prof}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}} \\ \mathcal{R}_{12} &= \{\text{Fellow}(\text{anna}), \text{Lect}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie})\} \cup \mathcal{A}_{\text{Int}}\end{aligned}$$

where the ABox \mathcal{A}_{Int} that is common to all the repairs is as follows:

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})\}$$

INCONSISTENCY-TOLERANT SEMANTICS: DEFINITIONS AND PROPERTIES

Repair: \subseteq -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

Repair: \subseteq -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

Plausible answers: hold no matter which repair is chosen

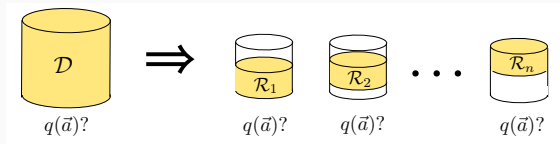
Repair: \subseteq -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

Plausible answers: hold no matter which repair is chosen

AR semantics: query each repair separately, intersect results

$$\mathcal{K} \models_{\text{AR}} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{B} \rangle \models q(\vec{a}) \text{ for every repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$



Reconsider our **example KB** $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

For the query $q_1(x) = \text{Fac}(x)$, we have:

Reconsider our **example KB** $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

For the query $q_1(x) = \text{Fac}(x)$, we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{anna})$, as every repair contains one of $\text{Prof}(\text{anna})$, $\text{Lect}(\text{anna})$, and $\text{Fellow}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{kim})$, as every repair contains $\text{Prof}(\text{kim})$ or $\text{Lect}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$, as every repair contains $\text{Fellow}(\text{alex})$

Reconsider our **example KB** $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

For the query $q_1(x) = \text{Fac}(x)$, we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{anna})$, as every repair contains one of $\text{Prof}(\text{anna})$, $\text{Lect}(\text{anna})$, and $\text{Fellow}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{kim})$, as every repair contains $\text{Prof}(\text{kim})$ or $\text{Lect}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$, as every repair contains $\text{Fellow}(\text{alex})$

These are the **only answers under AR semantics**:

EXAMPLE: AR SEMANTICS

Reconsider our **example KB** $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

For the query $q_1(x) = \text{Fac}(x)$, we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{anna})$, as every repair contains one of $\text{Prof}(\text{anna})$, $\text{Lect}(\text{anna})$, and $\text{Fellow}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{kim})$, as every repair contains $\text{Prof}(\text{kim})$ or $\text{Lect}(\text{anna})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$, as every repair contains $\text{Fellow}(\text{alex})$

These are the **only answers under AR semantics**:

- $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{julie})$ as $\langle \mathcal{T}_{\text{univ}}, \mathcal{R}_7 \rangle \not\models \text{Fac}(\text{julie})$
- can similarly show $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{CSC486})$ and $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{CSC343})$

Reconsider our **example KB** $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

For the query $q_2 = \exists y \text{Teaches}(x, y)$, we have:

Reconsider our **example KB** $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

For the query $q_2 = \exists y \text{Teaches}(x, y)$, we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{kim})$, as every repair contains $\text{Prof}(\text{kim})$ or $\text{Lect}(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{alex})$, as every repair contains $\text{Teaches}(\text{alex}, \text{csc486})$

Reconsider our **example KB** $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

For the query $q_2 = \exists y \text{Teaches}(x, y)$, we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{kim})$, as every repair contains $\text{Prof}(\text{kim})$ or $\text{Lect}(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_2(\text{alex})$, as every repair contains $\text{Teaches}(\text{alex}, \text{csc486})$

These are the **only answers under AR semantics**:

- $\mathcal{K}_{\text{univ}} \not\models_{\text{AR}} q_1(\text{anna})$ as $\langle \mathcal{T}_{\text{univ}}, \mathcal{R}_3 \rangle \not\models \exists y \text{Teaches}(\text{anna}, y)$
- can similarly show **julie**, **csc486**, and **csc343** are **not answers**

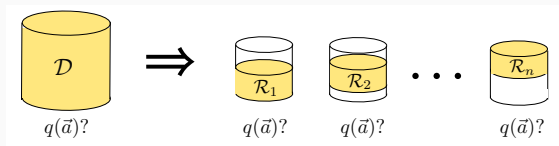
Repair: \subseteq -maximal subset of the data consistent with the ontology

- ways to achieve consistency, keeping as much information as possible

Plausible answers: hold no matter which repair is chosen

AR semantics: query each repair separately, intersect results

$$\mathcal{K} \models_{\text{AR}} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{B} \rangle \models q(\vec{a}) \text{ for every repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$



Satisfies both **CONSISTENT SUPPORT** and **CONSISTENT RESULTS**

Idea: only **use the surest assertions to answer queries**

- disregard assertions involved in some contradiction

Idea: only **use the surest assertions to answer queries**

- disregard assertions involved in some contradiction

IAR semantics: query the **intersection of the repairs**

$$\mathcal{K} \models_{\text{IAR}} q(\vec{a}) \quad \Leftrightarrow \quad \langle \mathcal{T}, \mathcal{D} \rangle \models q(\vec{a}) \text{ where } \mathcal{D} = \bigcap_{\mathcal{B} \in \text{Rep}(\mathcal{K})} \mathcal{B}$$

Reconsider our **example KB** $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

Intersection of the repairs of $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$:

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})\}$$

For the query $q_1(x) = \text{Fac}(x)$, we have:

Reconsider our **example KB** $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

Intersection of the repairs of $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$:

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})\}$$

For the query $q_1(x) = \text{Fac}(x)$, we have:

- $\mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex})$, as $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{Int}} \rangle \models \text{Fac}(\text{alex})$

EXAMPLE: IAR SEMANTICS

Reconsider our **example KB** $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$

Intersection of the repairs of $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$:

$$\mathcal{A}_{\text{Int}} = \{\text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486})\}$$

For the query $q_1(x) = \text{Fac}(x)$, we have:

$$\cdot \mathcal{K}_{\text{univ}} \models_{\text{AR}} q_1(\text{alex}), \text{ as } \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{Int}} \rangle \models \text{Fac}(\text{alex})$$

This is the **only answer to q_1** under **IAR semantics**:

- **anna** and **kim** are **no longer considered answers** since needed to **reason by cases** (e.g., kim is either Prof or Lect)

Idea: only **use the surest assertions to answer queries**

- disregard assertions involved in some contradiction

IAR semantics: query the **intersection of the repairs**

$$\mathcal{K} \models_{\text{IAR}} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{D} \rangle \models q(\vec{a}) \text{ where } \mathcal{D} = \bigcap_{\mathcal{B} \in \text{Rep}(\mathcal{K})} \mathcal{B}$$

Under-approximation of the **AR semantics**

Satisfies both **CONSISTENT SUPPORT** and **CONSISTENT RESULTS**

Idea: return all answers **supported by consistent part of data**

- can view them as **possible answers**, having **coherent justification**

Idea: return all answers **supported by consistent part of data**

- can view them as **possible answers**, having **coherent justification**

Brave semantics: query the repairs, take union of their answers

$$\mathcal{K} \models_{\text{brave}} q(\vec{a}) \iff \langle \mathcal{T}, \mathcal{B} \rangle \models q(\vec{a}) \text{ for some repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$

EXAMPLE: BRAVE SEMANTICS

Reconsider the KB $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

EXAMPLE: BRAVE SEMANTICS

Reconsider the KB $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

Moving from AR to brave semantics yields an **additional answer**:

- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{anna})$ AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{kim})$ AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{alex})$ AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{julie})$ $\langle \mathcal{T}_{\text{univ}}, \mathcal{R}_i \rangle \models q_1(\text{julie})$ for $1 \leq i \leq 6$

EXAMPLE: BRAVE SEMANTICS

Reconsider the KB $\mathcal{K}_{\text{univ}} = \langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

Moving from AR to brave semantics yields an **additional answer**:

- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{anna})$ AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{kim})$ AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{alex})$ AR-answer
- $\mathcal{K}_{\text{univ}} \models_{\text{brave}} q_1(\text{julie})$ $\langle \mathcal{T}_{\text{univ}}, \mathcal{R}_i \rangle \models q_1(\text{julie})$ for $1 \leq i \leq 6$

These are the **only answers to q_1 under brave semantics**:

- **csc486** and **csc343** cannot be obtained as answers from any repair

Idea: return all answers **supported by consistent part of data**

- can view them as **possible answers**, having **coherent justification**

Brave semantics: query the repairs, take union of their answers

$$\mathcal{K} \models_{\text{brave}} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{B} \rangle \models q(\vec{a}) \text{ for some repair } \mathcal{B} \in \text{Rep}(\mathcal{K})$$

Over-approximation of the **AR semantics**

- ... and **every semantics that satisfies CONSISTENT SUPPORT**

Does not satisfy CONSISTENT RESULTS

Why?

Goal: more fine-grained under-approximations of AR semantics

Goal: **more fine-grained under-approximations of AR semantics**

k -support semantics: bound # of supports used to 'cover' all repairs

$\mathcal{K} \models_{k\text{-supp}} q(\vec{a})$ iff exist subsets S_1, \dots, S_k of \mathcal{A} such that:

- each S_i is a \mathcal{T} -support for $q(\vec{a})$ in \mathcal{A}
- for every $R \in \text{Rep}(\mathcal{K})$, there is some S_i with $S_i \subseteq R$

Observe $\mathcal{K} \models_{k\text{-supp}} q(\vec{a}) \Rightarrow \mathcal{K} \models_{k+1\text{-supp}} q(\vec{a})$

EXAMPLE: K-SUPPORT SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

EXAMPLE: K-SUPPORT SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

When $k = 1$, the 1-support semantics gives **same result as IAR**:

- $\mathcal{K}_{\text{univ}} \models_{1\text{-supp}} q_1(\text{alex})$ use $\{\text{Fellow}(\text{alex})\}$

When $k = 2$, obtain **one additional answer**:

- $\mathcal{K}_{\text{univ}} \models_{2\text{-supp}} q_1(\text{alex})$
- $\mathcal{K}_{\text{univ}} \models_{2\text{-supp}} q_1(\text{kim})$ use $S_1 = \{\text{Prof}(\text{kim})\}$ and $S_2 = \{\text{Lect}(\text{kim})\}$

EXAMPLE: K-SUPPORT SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

When $k = 1$, the 1-support semantics gives **same result as IAR**:

- $\mathcal{K}_{\text{univ}} \models_{1\text{-supp}} q_1(\text{alex})$ use $\{\text{Fellow}(\text{alex})\}$

When $k = 2$, obtain **one additional answer**:

- $\mathcal{K}_{\text{univ}} \models_{2\text{-supp}} q_1(\text{alex})$
- $\mathcal{K}_{\text{univ}} \models_{2\text{-supp}} q_1(\text{kim})$ use $S_1 = \{\text{Prof}(\text{kim})\}$ and $S_2 = \{\text{Lect}(\text{kim})\}$

When $k = 3$, obtain **all three AR-answers**:

- $\mathcal{K}_{\text{univ}} \models_{2\text{-supp}} q_1(\text{alex})$
- $\mathcal{K}_{\text{univ}} \models_{2\text{-supp}} q_1(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{3\text{-supp}} q_1(\text{anna})$ use $S_1 = \{\text{Prof}(\text{anna})\}$, $S_2 = \{\text{Lect}(\text{anna})\}$,
and $S_3 = \{\text{Fellow}(\text{anna})\}$

Goal: **more fine-grained under-approximations of AR semantics**

k -support semantics: bound # of supports used to 'cover' all repairs

$\mathcal{K} \models_{k\text{-supp}} q(\vec{a})$ iff exist subsets S_1, \dots, S_k of \mathcal{A} such that:

- each S_i is a \mathcal{T} -support for $q(\vec{a})$ in \mathcal{A}
- for every $R \in \text{Rep}(\mathcal{K})$, there is some S_i with $S_i \subseteq R$

Observe $\mathcal{K} \models_{k+1\text{-supp}} q(\vec{a}) \Rightarrow \mathcal{K} \models_{k\text{-supp}} q(\vec{a})$

Relation to other semantics:

- when $k = 1$, same as IAR semantics
- for every $k \geq 1$, under-approximation of AR semantics
- for every \mathcal{K} , there is $k \geq 1$ such that $\mathcal{K} \models_{k\text{-supp}} q(\vec{a})$ iff $\mathcal{K} \models_{\text{AR}} q(\vec{a})$

Goal: **more fine-grained under-approximations of AR semantics**

k -support semantics: bound # of supports used to 'cover' all repairs

$\mathcal{K} \models_{k\text{-supp}} q(\vec{a})$ iff exist subsets S_1, \dots, S_k of \mathcal{A} such that:

- each S_i is a \mathcal{T} -support for $q(\vec{a})$ in \mathcal{A}
- for every $R \in \text{Rep}(\mathcal{K})$, there is some S_i with $S_i \subseteq R$

Observe $\mathcal{K} \models_{k+1\text{-supp}} q(\vec{a}) \Rightarrow \mathcal{K} \models_{k\text{-supp}} q(\vec{a})$

Relation to other semantics:

- when $k = 1$, same as IAR semantics
- for every $k \geq 1$, under-approximation of AR semantics
- for every \mathcal{K} , there is $k \geq 1$ such that $\mathcal{K} \models_{k\text{-supp}} q(\vec{a})$ iff $\mathcal{K} \models_{\text{AR}} q(\vec{a})$

Satisfy both **CONSISTENT SUPPORT** and **CONSISTENT RESULTS**

Goal: more fine-grained over-approximations of AR semantics

Goal: **more fine-grained over-approximations of AR semantics**

k -defeater semantics: bound # of assertions to block all supports

$\mathcal{K} \models_{k\text{-def}} q(\vec{a})$ iff does not exist a \mathcal{T} -consistent subset S of \mathcal{A} with:

- $|S| \leq k$
- $\langle \mathcal{T}, S \cup C \rangle \models \perp$ for every \subseteq -minimal \mathcal{T} -support $C \subseteq \mathcal{A}$ of $q(\vec{a})$

Observe $\mathcal{K} \models_{k+1\text{-def}} q(\vec{a}) \Rightarrow \mathcal{K} \models_{k\text{-def}} q(\vec{a})$

EXAMPLE: K-DEFEATER SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_2(x) = \exists y \text{Teaches}(x, y)$.

EXAMPLE: K-DEFEATER SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_2(x) = \exists y \text{Teaches}(x, y)$.

EXAMPLE: K-DEFEATER SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_2(x) = \exists y \text{Teaches}(x, y)$.

When $k = 0$, same answers as for **brave semantics**:

- `anna`, `kim`, `alex`, and `csc343`

EXAMPLE: K-DEFEATER SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_2(x) = \exists y \text{Teaches}(x, y)$.

When $k = 0$, same answers as for **brave semantics**:

- anna, kim, alex, and csc343

When $k = 1$, we **'lose' answers** anna and csc343:

- $\mathcal{K}_{\text{univ}} \not\models_{1\text{-def}} q_2(\text{anna})$ $\{\text{Fellow}(\text{anna})\}$ contradicts both minimal supports: $\{\text{Prof}(\text{anna})\}$ and $\{\text{Lect}(\text{anna})\}$
- $\mathcal{K}_{\text{univ}} \not\models_{1\text{-def}} q_2(\text{csc343})$ $\{\text{Fellow}(\text{julie})\}$ contradicts only minimal support $\{\text{Teaches}(\text{csc343}, \text{julie})\}$

EXAMPLE: K-DEFEATER SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_2(x) = \exists y \text{Teaches}(x, y)$.

When $k = 0$, same answers as for **brave semantics**:

- `anna`, `kim`, `alex`, and `csc343`

When $k = 1$, we **'lose' answers** `anna` and `csc343`:

- $\mathcal{K}_{\text{univ}} \not\models_{1\text{-def}} q_2(\text{anna})$ $\{\text{Fellow}(\text{anna})\}$ contradicts both minimal supports: $\{\text{Prof}(\text{anna})\}$ and $\{\text{Lect}(\text{anna})\}$
- $\mathcal{K}_{\text{univ}} \not\models_{1\text{-def}} q_2(\text{csc343})$ $\{\text{Fellow}(\text{julie})\}$ contradicts only minimal support $\{\text{Teaches}(\text{csc343}, \text{julie})\}$

Two other answers continue to hold under 1-support semantics:

- $\mathcal{K}_{\text{univ}} \models_{1\text{-def}} q_2(\text{kim})$
- $\mathcal{K}_{\text{univ}} \models_{1\text{-def}} q_2(\text{alex})$

Goal: **more fine-grained over-approximations of AR semantics**

k-defeater semantics: bound # of assertions to block all supports

$\mathcal{K} \models_{k\text{-def}} q(\vec{a})$ iff does not exist a \mathcal{T} -consistent subset S of \mathcal{A} with:

- $|S| \leq k$
- $\langle \mathcal{T}, S \cup C \rangle \models \perp$ for every \subseteq -minimal \mathcal{T} -support $C \subseteq \mathcal{A}$ of $q(\vec{a})$

Observe $\mathcal{K} \models_{k\text{-def}} q(\vec{a}) \Rightarrow \mathcal{K} \models_{k+1\text{-def}} q(\vec{a})$

Relation to other semantics:

- when $k = 0$, same as brave semantics
- for every $k \geq 0$, over-approximation of AR semantics
- for every \mathcal{K} , there is $k \geq 0$ such that $\mathcal{K} \models_{k\text{-def}} q(\vec{a})$ iff $\mathcal{K} \models_{\text{AR}} q(\vec{a})$

Goal: **more fine-grained over-approximations of AR semantics**

k-defeater semantics: bound # of assertions to block all supports

$\mathcal{K} \models_{k\text{-def}} q(\vec{a})$ iff does not exist a \mathcal{T} -consistent subset S of \mathcal{A} with:

- $|S| \leq k$
- $\langle \mathcal{T}, S \cup C \rangle \models \perp$ for every \subseteq -minimal \mathcal{T} -support $C \subseteq \mathcal{A}$ of $q(\vec{a})$

Observe $\mathcal{K} \models_{k\text{-def}} q(\vec{a}) \Rightarrow \mathcal{K} \models_{k+1\text{-def}} q(\vec{a})$

Relation to other semantics:

- when $k = 0$, same as brave semantics
- for every $k \geq 0$, over-approximation of AR semantics
- for every \mathcal{K} , there is $k \geq 0$ such that $\mathcal{K} \models_{k\text{-def}} q(\vec{a})$ iff $\mathcal{K} \models_{\text{AR}} q(\vec{a})$

Satisfy CONSISTENT SUPPORT but not **CONSISTENT RESULTS**

Goal: obtain **closer under-approximation of AR than IAR**

$\text{close}_{\mathcal{T}}(\mathcal{A}) = \text{all ABox assertions entailed from } \langle \mathcal{T}, \mathcal{A} \rangle$

ICR semantics: close repairs, intersect them, then query the result

$\mathcal{K} \models_{\text{ICR}} q(\vec{a}) \text{ iff } \langle \mathcal{T}, \mathcal{D} \rangle \models q(\vec{a}) \text{ where } \mathcal{D} = \bigcap_{\mathcal{B} \in \text{Rep}(\mathcal{K})} \text{close}_{\mathcal{T}}(\mathcal{B})$

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

EXAMPLE: ICR SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

Close the repairs of $\mathcal{K}_{\text{univ}}$:

$$\text{close}_{\mathcal{T}_{\text{univ}}}(\mathcal{R}_1) = \{\text{Prof}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fac}(\text{anna}), \\ \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Fac}(\text{julie}), \text{Course}(\text{csc486})\} \\ \vdots$$

$$\text{close}_{\mathcal{T}_{\text{univ}}}(\mathcal{R}_{12}) = \{\text{Fellow}(\text{anna}), \text{Lect}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie}), \\ \text{Fac}(\text{anna}), \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Fac}(\text{julie}), \text{Course}(\text{julie})\}$$

EXAMPLE: ICR SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

Close the repairs of $\mathcal{K}_{\text{univ}}$:

$$\text{close}_{\mathcal{T}_{\text{univ}}}(\mathcal{R}_1) = \{\text{Prof}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fac}(\text{anna}), \\ \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Fac}(\text{julie}), \text{Course}(\text{csc486})\} \\ \vdots$$

$$\text{close}_{\mathcal{T}_{\text{univ}}}(\mathcal{R}_{12}) = \{\text{Fellow}(\text{anna}), \text{Lect}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie}), \\ \text{Fac}(\text{anna}), \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Fac}(\text{julie}), \text{Course}(\text{julie})\}$$

Take intersection of the closed repairs:

$$\mathcal{A}'_{\text{int}} = \{\text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486}), \text{Fac}(\text{anna}), \\ \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Course}(\text{csc486})\}$$

EXAMPLE: ICR SEMANTICS

Reconsider the KB $\langle \mathcal{T}_{\text{univ}}, \mathcal{A}_{\text{univ}} \rangle$ and query $q_1(x) = \text{Fac}(x)$.

Close the repairs of $\mathcal{K}_{\text{univ}}$:

$$\begin{aligned} \text{close}_{\mathcal{T}_{\text{univ}}}(\mathcal{R}_1) &= \{\text{Prof}(\text{anna}), \text{Prof}(\text{kim}), \text{Fellow}(\text{julie}), \text{Fac}(\text{anna}), \\ &\quad \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Fac}(\text{julie}), \text{Course}(\text{csc486})\} \\ &\quad \vdots \end{aligned}$$

$$\begin{aligned} \text{close}_{\mathcal{T}_{\text{univ}}}(\mathcal{R}_{12}) &= \{\text{Fellow}(\text{anna}), \text{Lect}(\text{kim}), \text{Teaches}(\text{csc343}, \text{julie}), \\ &\quad \text{Fac}(\text{anna}), \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Fac}(\text{julie}), \text{Course}(\text{julie})\} \end{aligned}$$

Take intersection of the closed repairs:

$$\begin{aligned} \mathcal{A}'_{\text{int}} &= \{\text{Fellow}(\text{alex}), \text{Teaches}(\text{alex}, \text{csc486}), \text{Fac}(\text{anna}), \\ &\quad \text{Fac}(\text{kim}), \text{Fac}(\text{alex}), \text{Course}(\text{csc486})\} \end{aligned}$$

Get following ICR-answers: anna, kim, alex

Goal: obtain **closer under-approximation of AR than IAR**

$\text{close}_{\mathcal{T}}(\mathcal{A}) = \text{all ABox assertions entailed from } \langle \mathcal{T}, \mathcal{A} \rangle$

ICR semantics: close repairs, intersect them, then query the result

$\mathcal{K} \models_{\text{ICR}} q(\vec{a}) \text{ iff } \langle \mathcal{T}, \mathcal{D} \rangle \models q(\vec{a}) \text{ where } \mathcal{D} = \bigcap_{\mathcal{B} \in \text{Rep}(\mathcal{K})} \text{close}_{\mathcal{T}}(\mathcal{B})$

Under-approximation of AR semantics, over-approximation of IAR

- same as AR semantics for IQs and quantifier-free CQs

Goal: obtain **closer under-approximation of AR than IAR**

$\text{close}_{\mathcal{T}}(\mathcal{A})$ = all **ABox assertions entailed from $\langle \mathcal{T}, \mathcal{A} \rangle$**

ICR semantics: close repairs, intersect them, then query the result

$\mathcal{K} \models_{\text{ICR}} q(\vec{a})$ iff $\langle \mathcal{T}, \mathcal{D} \rangle \models q(\vec{a})$ where $\mathcal{D} = \bigcap_{\mathcal{B} \in \text{Rep}(\mathcal{K})} \text{close}_{\mathcal{T}}(\mathcal{B})$

Under-approximation of AR semantics, over-approximation of IAR

· same as AR semantics for IQs and quantifier-free CQs

Satisfies CONSISTENT SUPPORT and CONSISTENT RESULTS

Goal: define semantics that are (almost) **syntax-independent**

- apply **closure** operator on **original ABox**

Goal: define semantics that are (almost) **syntax-independent**

- apply **closure** operator on **original ABox**

Need **alternative notion of closure**:

$$\text{close}_{\mathcal{T}}^*(\mathcal{A}) = \{\beta \mid \exists S \subseteq \mathcal{A} \text{ such that } S \text{ is } \mathcal{T}\text{-consistent and } \langle \mathcal{T}, S \rangle \models \beta\}$$

Goal: define semantics that are (almost) **syntax-independent**

- apply **closure** operator on **original ABox**

Need **alternative notion of closure**:

$$\text{close}_{\mathcal{T}}^*(\mathcal{A}) = \{\beta \mid \exists S \subseteq \mathcal{A} \text{ such that } S \text{ is } \mathcal{T}\text{-consistent and } \langle \mathcal{T}, S \rangle \models \beta\}$$

Closed ABox repair: maximally **'complete'** the **standard ABox repairs** with additional **facts from** $\text{close}_{\mathcal{T}}^*(\mathcal{A}) \setminus \mathcal{A}$

Goal: define semantics that are (almost) **syntax-independent**

- apply **closure** operator on **original ABox**

Need **alternative notion of closure**:

$$\text{close}_{\mathcal{T}}^*(\mathcal{A}) = \{\beta \mid \exists S \subseteq \mathcal{A} \text{ such that } S \text{ is } \mathcal{T}\text{-consistent and } \langle \mathcal{T}, S \rangle \models \beta\}$$

Closed ABox repair: maximally **'complete'** the **standard ABox repairs** with additional **facts from** $\text{close}_{\mathcal{T}}^*(\mathcal{A}) \setminus \mathcal{A}$

CAR semantics = **AR** semantics but using **closed ABox repairs**

ICAR semantics = **IAR** semantics but using **closed ABox repairs**

Relations with other semantics:

- **ICAR** semantics is an **under-approximation** of the **CAR** semantics
- **CAR** semantics is an **over-approximation** of the **AR** semantics
- **ICAR** semantics is an **over-approximation** of the **ICR** semantics
- **CAR and ICAR** are **not under-approximations** of **brave** semantics

Relations with other semantics:

- **ICAR** semantics is an **under-approximation** of the **CAR** semantics
- **CAR** semantics is an **over-approximation** of the **AR** semantics
- **ICAR** semantics is an **over-approximation** of the **ICR** semantics
- **CAR and ICAR** are **not under-approximations** of **brave** semantics

CAR and ICAR semantics **satisfy CONSISTENT RESULTS**

Relations with other semantics:

- **ICAR** semantics is an **under-approximation** of the **CAR** semantics
- **CAR** semantics is an **over-approximation** of the **AR** semantics
- **ICAR** semantics is an **over-approximation** of the **ICR** semantics
- **CAR and ICAR** are **not under-approximations** of **brave** semantics

CAR and ICAR semantics **satisfy CONSISTENT RESULTS**

These semantics **do not satisfy CONSISTENT SUPPORT**

$$\mathcal{T} = \{A \sqsubseteq B, C \sqsubseteq D, A \sqsubseteq \neg C\}, \mathcal{A} = \{A(e), C(e)\}, \text{ and } q = B(x) \wedge D(x)$$

Clusters: partition of ABox assertions

- **group together assertions** that **appear together** in **minimal \mathcal{T} -inconsistent subset**

Clusters: partition of ABox assertions

- group together assertions that appear together in minimal \mathcal{T} -inconsistent subset

k-lazy repairs: for each cluster C_i , remove from \mathcal{A} either

- minimal subset $C'_i \subseteq C_i$ such that $|C'_i| \leq k$ and $C_i \setminus C'_i$ is \mathcal{T} -consistent
- if no such C'_i exists, remove the whole cluster C_i

Clusters: partition of ABox assertions

- group together assertions that appear together in minimal \mathcal{T} -inconsistent subset

k-lazy repairs: for each cluster C_i , remove from \mathcal{A} either

- minimal subset $C'_i \subseteq C_i$ such that $|C'_i| \leq k$ and $C_i \setminus C'_i$ is \mathcal{T} -consistent
- if no such C'_i exists, remove the whole cluster C_i

k-lazy semantics: like AR, but use k-lazy repairs

Relation to other semantics:

- $\mathcal{K} \models_{\text{IAR}} q(\vec{a})$ iff $\mathcal{K} \models_{0\text{-lazy}} q(\vec{a})$
- for every $k \geq 0$, if $\mathcal{K} \models_{k\text{-lazy}} q(\vec{a})$, then $\mathcal{K} \models_{\text{brave}} q(\vec{a})$
- for **every** KB \mathcal{K} , there **exists some** $k \geq 0$ such that for **every** $k' \geq k$:
 $\mathcal{K} \models_{\text{AR}} q(\vec{a})$ iff $\mathcal{K} \models_{k'\text{-lazy}} q(\vec{a})$

Relation to other semantics:

- $\mathcal{K} \models_{\text{IAR}} q(\vec{a})$ iff $\mathcal{K} \models_{0\text{-lazy}} q(\vec{a})$
- for every $k \geq 0$, if $\mathcal{K} \models_{k\text{-lazy}} q(\vec{a})$, then $\mathcal{K} \models_{\text{brave}} q(\vec{a})$
- for **every** KB \mathcal{K} , there **exists some** $k \geq 0$ such that for **every** $k' \geq k$:
 $\mathcal{K} \models_{\text{AR}} q(\vec{a})$ iff $\mathcal{K} \models_{k'\text{-lazy}} q(\vec{a})$

Convergence not monotone in k :

- possible to have $\mathcal{K} \models_{k\text{-lazy}} q(\vec{a})$ but $\mathcal{K} \not\models_{k+1\text{-lazy}} q(\vec{a})$

Relation to other semantics:

- $\mathcal{K} \models_{\text{IAR}} q(\vec{a})$ iff $\mathcal{K} \models_{0\text{-lazy}} q(\vec{a})$
- for every $k \geq 0$, if $\mathcal{K} \models_{k\text{-lazy}} q(\vec{a})$, then $\mathcal{K} \models_{\text{brave}} q(\vec{a})$
- for **every** KB \mathcal{K} , there **exists some** $k \geq 0$ such that for **every** $k' \geq k$:
 $\mathcal{K} \models_{\text{AR}} q(\vec{a})$ iff $\mathcal{K} \models_{k'\text{-lazy}} q(\vec{a})$

Convergence not monotone in k :

- possible to have $\mathcal{K} \models_{k\text{-lazy}} q(\vec{a})$ but $\mathcal{K} \not\models_{k+1\text{-lazy}} q(\vec{a})$

CONSISTENT SUPPORT and **CONSISTENT RESULTS** satisfied (for every k)

Idea: **some repairs** are **more likely than others**

- exploit knowledge about relative **reliability of ABox assertions**

Idea: **some repairs** are **more likely than others**

- exploit knowledge about relative **reliability of ABox assertions**

Formally: use **preference relation** \preceq to **compare repairs**

- compare w.r.t. **cardinality** (\leq)
- partition ABox into **priority levels** $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$
 - compare **level-by-level using set inclusion** ($\subseteq_{\mathcal{P}}$)
 - compare **level-by-level using cardinality** ($\leq_{\mathcal{P}}$)
- assign **weights to ABox assertions**
 - compare repairs by **total weight** (\leq_w)

Idea: **some repairs** are **more likely than others**

- exploit knowledge about relative **reliability of ABox assertions**

Formally: use **preference relation** \preceq to **compare repairs**

- compare w.r.t. **cardinality** (\leq)
- partition ABox into **priority levels** $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$
 - compare **level-by-level using set inclusion** ($\subseteq_{\mathcal{P}}$)
 - compare **level-by-level using cardinality** ($\leq_{\mathcal{P}}$)
- assign **weights to ABox assertions**
 - compare repairs by **total weight** (\leq_w)

AR / IAR / brave semantics based upon **most preferred repairs**
(\preceq -AR, \preceq -IAR, \preceq -brave)

COMPLEXITY OF INCONSISTENCY- TOLERANT QUERY ANSWERING

View OMQA as a **decision problem** (yes-or-no question):

- PROBLEM: **\mathcal{Q} answering in \mathcal{L} under semantics S**
(\mathcal{Q} a query language, \mathcal{L} a DL, S chosen semantics)
- INPUT: An **n -ary query $q \in \mathcal{Q}$** , an **ABox \mathcal{A}** , a **\mathcal{L} -TBox \mathcal{T}** ,
and a **tuple $\vec{a} \in \text{Ind}(\mathcal{A})^n$**
- QUESTION: **Does $\langle \mathcal{T}, \mathcal{A} \rangle \models_S q(\vec{a})$?**

View OMQA as a **decision problem** (yes-or-no question):

PROBLEM: \mathcal{Q} answering in \mathcal{L} under semantics S
(\mathcal{Q} a query language, \mathcal{L} a DL, S chosen semantics)

INPUT: An n -ary query $q \in \mathcal{Q}$, an **ABox** \mathcal{A} , a \mathcal{L} -TBox \mathcal{T} ,
and a **tuple** $\vec{a} \in \text{Ind}(\mathcal{A})^n$

QUESTION: **Does** $\langle \mathcal{T}, \mathcal{A} \rangle \models_S q(\vec{a})$?

Combined complexity: in terms of **size of whole input**

Data complexity: in terms of **size of \mathcal{A} only**

- **view rest of input as fixed** (of constant size)
- **motivation: ABox typically much larger than rest of input**

Note: use $|\mathcal{A}|$ to denote **size of \mathcal{A}** (similarly for $|\mathcal{T}|$, $|q|$, etc.)

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite \mathcal{R}** and **all DL-Lite dialects that satisfy:**

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite \mathcal{R}** and **all DL-Lite dialects that satisfy:**

- every **minimal support for $q(\vec{a})$** contains **at most $|q|$ assertions**

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite \mathcal{R}** and **all DL-Lite dialects that satisfy:**

- every **minimal support for $q(\vec{a})$** contains **at most $|q|$ assertions**
- every **minimal \mathcal{T} -inconsistent subset** has **cardinality at most two**

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite \mathcal{R}** and **all DL-Lite dialects that satisfy:**

- every **minimal support for $q(\vec{a})$** contains **at most $|q|$ assertions**
- every **minimal \mathcal{T} -inconsistent subset** has **cardinality at most two**
- **CQ answering**, IQ answering, and KB consistency can be performed by **FO query rewriting** (so in **AC⁰ in data complexity**)

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite \mathcal{R}** and **all DL-Lite dialects that satisfy:**

- every **minimal support for $q(\vec{a})$** contains **at most $|q|$ assertions**
- every **minimal \mathcal{T} -inconsistent subset** has **cardinality at most two**
- **CQ answering**, IQ answering, and KB consistency can be performed by **FO query rewriting** (so in **AC 0 in data complexity**)
- **CQ answering** is **NP-complete for combined complexity**

Today: mainly **focus on DL-Lite**, important DL for OMQA

Results apply to **DL-Lite \mathcal{R}** and **all DL-Lite dialects that satisfy:**

- every **minimal support for $q(\vec{a})$** contains **at most $|q|$ assertions**
- every **minimal \mathcal{T} -inconsistent subset** has **cardinality at most two**
- **CQ answering**, IQ answering, and KB consistency can be performed by **FO query rewriting** (so in **AC⁰ in data complexity**)
- **CQ answering** is **NP-complete for combined complexity**
- **IQ answering** is **NL-complete in combined complexity**

Theorem CQ and IQ answering under AR semantics are
coNP-complete in data complexity

Theorem CQ and IQ answering under AR semantics are
coNP-complete in data complexity

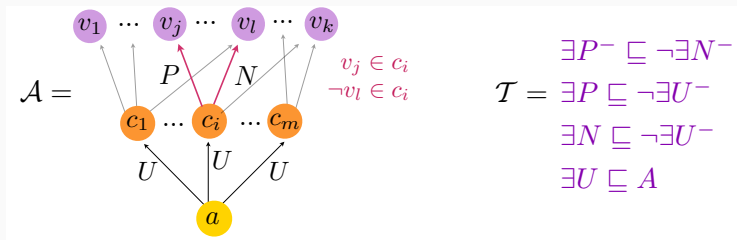
Upper bound: guess $\mathcal{A}' \subseteq \mathcal{A}$, verify \mathcal{A}' is repair and $\langle \mathcal{T}, \mathcal{A}' \rangle \not\models q(\vec{a})$

BAD NEWS: INTRACTABILITY OF AR SEMANTICS

Theorem CQ and IQ answering under AR semantics are
coNP-complete in data complexity

Upper bound: guess $\mathcal{A}' \subseteq \mathcal{A}$, verify \mathcal{A}' is repair and $\langle \mathcal{T}, \mathcal{A}' \rangle \models q(\vec{a})$

Lower bound: reduction from UNSAT $\varphi = c_1 \wedge \dots \wedge c_m$ over v_1, \dots, v_k



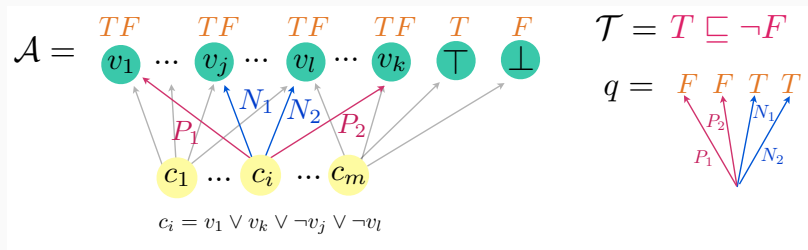
Can show φ unsatisfiable $\Leftrightarrow \mathcal{T}, \mathcal{A} \models_{AR} A(a)$

In fact: **CQ answering** is **coNP-hard** for simple TBox $\mathcal{T} = \{T \sqsubseteq \neg F\}$

BAD NEWS: INTRACTABILITY OF AR SEMANTICS (CONT.)

In fact: **CQ answering** is **coNP-hard** for simple TBox $\mathcal{T} = \{T \sqsubseteq \neg F\}$

Reduction from 2+2UNSAT: $\varphi = c_1 \wedge \dots \wedge c_m$ over $v_1, \dots, v_k, \top, \perp$
 each clause has **two positive** and **two negative literals**



Can show φ **unsatisfiable** $\Leftrightarrow \mathcal{T}, \mathcal{A} \models q$

Can use preceding reductions to show more intractability results

Can use preceding reductions to show more intractability results

Theorem **CQ and IQ answering** under **ICR semantics** are **coNP-complete** in **data complexity**

Theorem **CQ answering** under **CAR semantics** is **coNP-complete** in **data complexity**

Theorem **CQ answering** under **k-lazy semantics** is **coNP-complete** in **data complexity** for every $k \geq 1$

For IAR and brave semantics,
have **same low data complexity as classical semantics**

Theorem **CQ and IQ answering** under **IAR semantics** are
in AC^0 in **data complexity**

Theorem **CQ and IQ answering** under **brave semantics** are
in AC^0 in **data complexity**

For IAR and brave semantics,
have **same low data complexity as classical semantics**

Theorem **CQ and IQ answering** under **IAR semantics** are
in AC^0 in **data complexity**

Theorem **CQ and IQ answering** under **brave semantics** are
in AC^0 in **data complexity**

Can use **FO-query rewriting** to **compute IAR- and brave-answers**

Idea: **modify UCQ-rewriting** to ensure **ABox assertions** matching disjuncts are **not involved in any contradictions**

Idea: **modify UCQ-rewriting** to ensure **ABox assertions** matching disjuncts are **not involved in any contradictions**

(Normal) rewriting of $q_2(x) = \exists y \text{Teaches}(x, y)$ w.r.t. $\mathcal{T}_{\text{univ}}$:

$$q'_2(x) = \text{Prof}(x) \vee \text{Lect}(x) \vee \exists y. \text{Teaches}(x, y)$$

Idea: **modify UCQ-rewriting** to ensure **ABox assertions** matching disjuncts are **not involved in any contradictions**

(Normal) rewriting of $q_2(x) = \exists y \text{Teaches}(x, y)$ w.r.t. $\mathcal{T}_{\text{univ}}$:

$$q'_2(x) = \text{Prof}(x) \vee \text{Lect}(x) \vee \exists y. \text{Teaches}(x, y)$$

Rewriting of q_2 for **IAR semantics**:

$$\begin{aligned} q''_2(x) = & \text{Prof}(x) \wedge (\neg \text{Lect}(x) \wedge \neg \text{Fellow}(x) \wedge \neg \text{Course}(x) \wedge \neg \exists z. \text{Teaches}(z, x)) \vee \\ & \text{Lect}(x) \wedge (\neg \text{Prof}(x) \wedge \neg \text{Fellow}(x) \wedge \neg \text{Course}(x) \wedge \neg \exists z. \text{Teaches}(z, x)) \vee \\ & \exists y. (\text{Teaches}(x, y) \wedge (\neg \text{Prof}(y) \wedge \neg \text{Lect}(y) \wedge \neg \text{Fellow}(y))) \end{aligned}$$

Idea: **modify UCQ-rewriting** to ensure each **disjunct can only match \mathcal{T} -consistent subset of ABox**

Idea: **modify UCQ-rewriting** to ensure each **disjunct can only match**
 \mathcal{T} -consistent subset of ABox

Modified **TBox** $\mathcal{T}'_{\text{univ}}$: add $\exists \text{Teaches} \sqsubseteq \text{Fac}$ to $\mathcal{T}_{\text{univ}}$

Idea: **modify UCQ-rewriting** to ensure each **disjunct can only match \mathcal{T} -consistent subset of ABox**

Modified **TBox** $\mathcal{T}'_{\text{univ}}$: add $\exists \text{Teaches} \sqsubseteq \text{Fac}$ to $\mathcal{T}_{\text{univ}}$

(Normal) rewriting of $q_2(x) = \exists y \text{Teaches}(x, y)$ w.r.t. $\mathcal{T}'_{\text{univ}}$:

$$q'_2(x) = \text{Prof}(x) \vee \text{Lect}(x) \vee \exists y. \text{Teaches}(x, y)$$

Idea: **modify UCQ-rewriting** to ensure each **disjunct can only match** \mathcal{T} -consistent subset of ABox

Modified **TBox** $\mathcal{T}'_{\text{univ}}$: add $\exists \text{Teaches} \sqsubseteq \text{Fac}$ to $\mathcal{T}_{\text{univ}}$

(Normal) rewriting of $q_2(x) = \exists y \text{Teaches}(x, y)$ w.r.t. $\mathcal{T}'_{\text{univ}}$:

$$q'_2(x) = \text{Prof}(x) \vee \text{Lect}(x) \vee \exists y. \text{Teaches}(x, y)$$

Rewriting of q_2 for **brave semantics**:

$$q'_2(x) = \text{Prof}(x) \vee \text{Lect}(x) \vee (\exists y. \text{Teaches}(x, y) \wedge x \neq y)$$

to **disallow** using assertions of the form $\text{Teaches}(a, a)$

Theorem CQ and IQ answering under k -support semantics ($k \geq 1$) or k -defeater semantics ($k \geq 0$) are in AC^0 in data complexity

Theorem **CQ and IQ answering** under **k -support semantics** ($k \geq 1$) or **k -defeater semantics** ($k \geq 0$) are in AC^0 in **data complexity**

Shape of rewriting for k -support semantics:

$$\bigvee_{\substack{\text{choice of } k \\ \text{support-types}}} \left(\begin{array}{l} \text{chosen supports} \\ \text{present in the ABox} \end{array} \wedge \begin{array}{l} \text{supports cannot be} \\ \text{simultaneously contradicted} \end{array} \right)$$

Theorem **CQ and IQ answering** under **k -support semantics** ($k \geq 1$) or **k -defeater semantics** ($k \geq 0$) are in AC^0 in **data complexity**

Shape of rewriting for k -support semantics:

$$\bigvee_{\text{choice of } k \text{ support-types}} \left(\begin{array}{l} \text{chosen supports} \\ \text{present in the ABox} \end{array} \wedge \begin{array}{l} \text{supports cannot be} \\ \text{simultaneously contradicted} \end{array} \right)$$

Shape of rewriting for k -defeater semantics:

$$\neg \left(\bigvee_{\text{choice of defeater of size at most } k} \left(\begin{array}{l} \text{chosen defeater} \\ \text{present in the ABox} \end{array} \wedge \begin{array}{l} \text{contradicts every support} \\ \text{present in the ABox} \end{array} \right) \right)$$

Theorem **CQ and IQ answering** under **k -support semantics** ($k \geq 1$) or **k -defeater semantics** ($k \geq 0$) are in AC^0 in **data complexity**

Shape of rewriting for k -support semantics:

$$\bigvee_{\text{choice of } k \text{ support-types}} \left(\begin{array}{l} \text{chosen supports} \\ \text{present in the ABox} \end{array} \wedge \begin{array}{l} \text{supports cannot be} \\ \text{simultaneously contradicted} \end{array} \right)$$

Shape of rewriting for k -defeater semantics:

$$\neg \left(\bigvee_{\text{choice of defeater of size at most } k} \left(\begin{array}{l} \text{chosen defeater} \\ \text{present in the ABox} \end{array} \wedge \begin{array}{l} \text{contradicts every support} \\ \text{present in the ABox} \end{array} \right) \right)$$

Note: positive results hold for **all FO-rewritable ontology languages**

COMPLEXITY LANDSCAPE FOR DL-LITE

Semantics	Data complexity		Combined complexity	
	CQs	IQs	CQs	IQs
classical	in AC^0	in AC^0	NP	NL
AR	coNP	coNP	Π_2^P	coNP
IAR	in AC^0	in AC^0	NP	NL
brave	in AC^0	in AC^0	NP	NL
k -support ($k \geq 1$)	in AC^0	in AC^0	NP	NL
k -defeater ($k \geq 0$)	in AC^0	in AC^0	NP	NL
ICR	coNP	coNP	$\Delta_2^P[O(\log n)]$	coNP
CAR	coNP	in AC^0	Π_2^P	NL
ICAR	in AC^0	in AC^0	NP	NL
k -lazy ($k \geq 1$)	coNP	in P	Π_2^P	in P

CQAPri first system for AR query answering in DL-Lite

Implements **hybrid approach**:

- compute IAR and brave answers polytime
 - gives upper and lower **bounds on AR answers**
- use SAT solvers to **identify remaining AR answers**
- three categories of answers: **possible, likely, (almost) sure**

CQAPri first system for AR query answering in DL-Lite

Implements **hybrid approach**:

- compute IAR and brave answers polytime
 - gives upper and lower **bounds on AR answers**
- use SAT solvers to **identify remaining AR answers**
- three categories of answers: **possible, likely, (almost) sure**

Encouraging experimental results:

- in most cases, IAR and brave enough to decide if tuple is AR-answer \Rightarrow **few calls to SAT solvers**
- **SAT encodings** are typically **small and easy to solve**

Lightweight DL \mathcal{EL}_{\perp} : constructors $\top, \perp, \sqcap, \exists r.C$

Semantics	Data complexity		Combined complexity	
	CQs	IQs	CQs	IQs
classical	P	P	NP	P
AR	coNP	coNP	Π_2^P	coNP
IAR	coNP	coNP	$\Delta_2^P[O(\log n)]$	coNP
brave	NP	NP	NP	NP

Observe: **IAR** and **brave** are **no longer tractable**

- **no bound** on size of **minimal \mathcal{T} -inconsistent subsets**

Expressive DL \mathcal{ALC} : constructors $\top, \perp, \neg, \sqcap, \sqcup, \exists r.C, \forall r.C$

Semantics	Data complexity		Combined complexity	
	CQs	IQs	CQs	IQs
classical	coNP	coNP	EXP	EXP
AR	Π_2^P	Π_2^P	EXP	EXP
IAR	Π_2^P	Π_2^P	EXP	EXP
brave	Σ_2^P	Σ_2^P	EXP	EXP

Observe:

- IAR and brave no easier than AR
- increased data complexity, no increase in combined complexity

CONCLUSION & OUTLOOK

Traditional OMQA techniques **not robust to data inconsistencies**
⇒ **need for inconsistency-tolerant semantics**

Many different semantics have been proposed

- some **borrowed from other areas** (DBs, KR)
- many proposed **specifically for OMQA**

Good understanding of complexity landscape for inconsistency-tolerant OMQA

- **complete picture for DL-Lite**
- quite a few results for **other DLs, existential rules**

First implemented systems, promising results for DL-Lite

Algorithms for DL-Lite:

- mostly based upon UCQ-rewritings (can be very large)
- can we **adapt other techniques** to our setting?

Beyond DL-Lite:

- complexity landscape already explored, mostly negative results
- challenge: **how to design effective algorithms?**
- difficulty: **unbounded size of query supports** and **minimal inconsistent subsets**

Usability issues:

- need for **explanation** services: **help users interpret query results**
- try to **improve data quality** through **interaction with users**