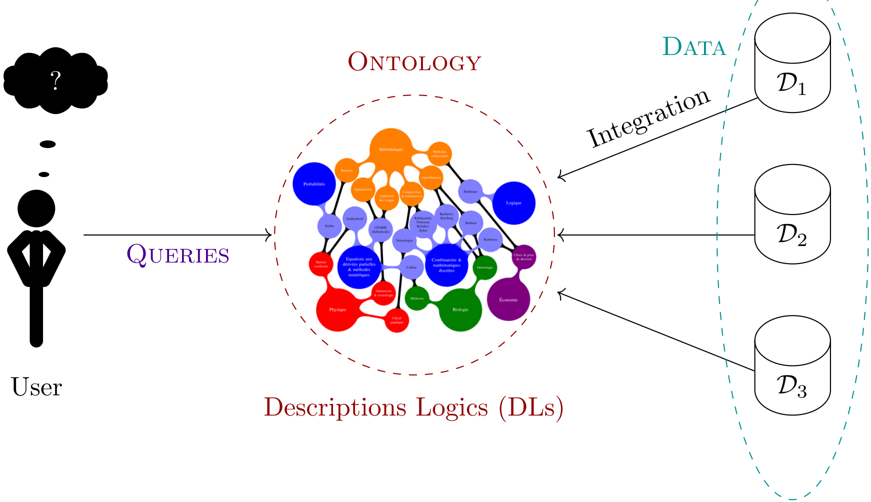
Answering Counting Queries over DL-Lite Ontologies

Meghyn Bienvenu[1], Quentin Manière[1], Michaël Thomazo[2]

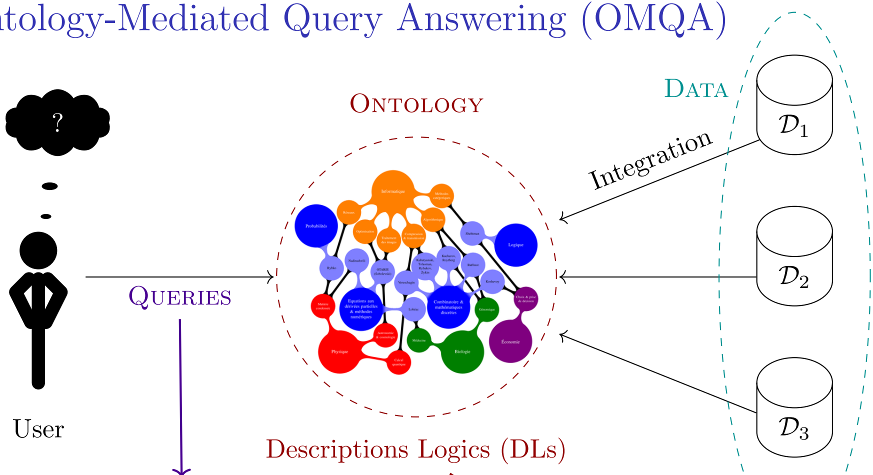
[1] University of Bordeaux, CNRS, Bordeaux INP, LaBRI, Talence, France
[2] Inria, DI ENS, ENS, CNRS, University PSL, Paris, France

Journée de la Logique, Montpellier

Ontology-Mediated Query Answering (OMQA)



Ontology-Mediated Query Answering (OMQA)



Descriptions Logics (DLs)

Conjunctive query (CQ):
how to satisfy a condition?

Counting CQ:
how many ways to satisfy a condition?

DL-Lite$_{\mathsf{core}}$, DL-Lite$_{\mathcal{R}}$

Example of an OMQA setting

TBox (ontology): Terminological axioms

$$\forall x \ (\ \text{Friend}(x) \ \rightarrow \ \exists y \ \text{listensTo}(x, y)\ ) \qquad \text{(FOL notation)}$$

$$\text{Friend} \ \sqsubseteq \ \exists \text{listensTo} \qquad \text{(DLs notation)}$$

Counting conjunctive query (CCQ)

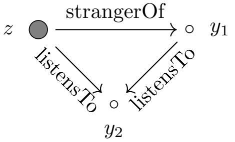
$$q := \exists \mathbf{y} \ \exists \mathbf{z} \ \psi(\mathbf{y}, \mathbf{z})$$

ψ: conjunction of concept or role atoms involving constants or variables

$\mathbf{y}$: tuple of *existential* variables

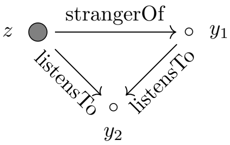
$\mathbf{z}$: tuple of *counting* variables

$$\exists y_1 \ \exists y_2 \ \exists z \ \text{strangerOf}(z, y_1) \wedge \text{listensTo}(z, y_2) \wedge \text{listensTo}(y_1, y_2)$$



How many friends z can I introduce to a friend y_1 based on a shared musical taste y_2?

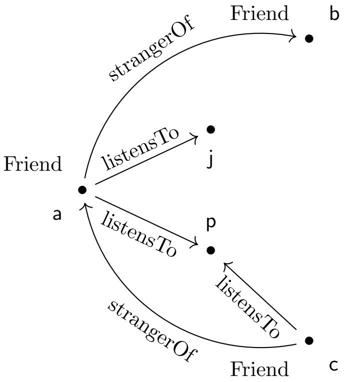
Example of an OMQA setting



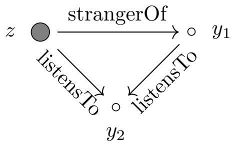
$$\forall x \; (\; \mathrm{Friend}(x) \; \rightarrow \; \exists y \; \mathrm{listensTo}(x, y) \;)$$

ABox (data): Assertions concerning constants

Friend(alex)
Friend(bill)
Friend(carol)
listensTo(alex, jazz)
listensTo(alex, pop)
listensTo(carol, pop)
strangerOf(alex, bill)
strangerOf(carol, alex)

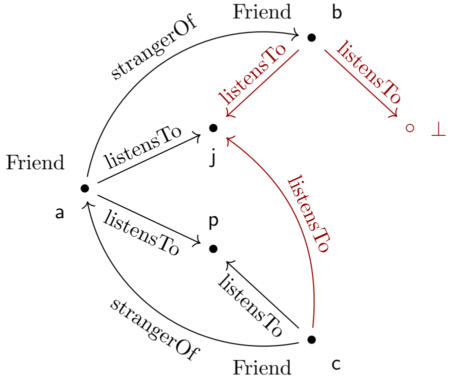


Example of an OMQA setting

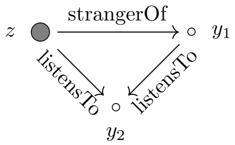


$\forall x \; (\; \text{Friend}(x) \; \rightarrow \; \exists y \; \text{listensTo}(x, y) \;)$

Model 1



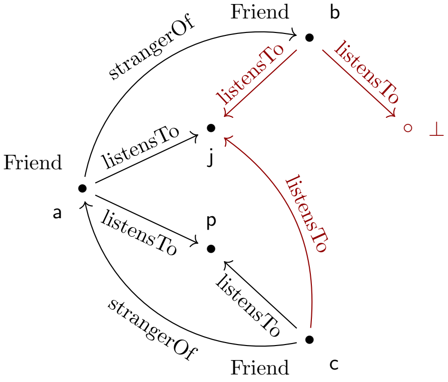
Example of an OMQA setting



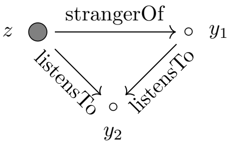
$$\forall x \; (\; \text{Friend}(x) \; \rightarrow \; \exists y \; \text{listensTo}(x, y) \;)$$

Model 1

$$\begin{array}{ccc} z & \mapsto & \mathbf{a} \\ y_1 & \mapsto & \mathsf{b} \\ y_2 & \mapsto & \mathsf{j} \end{array}$$



Example of an OMQA setting

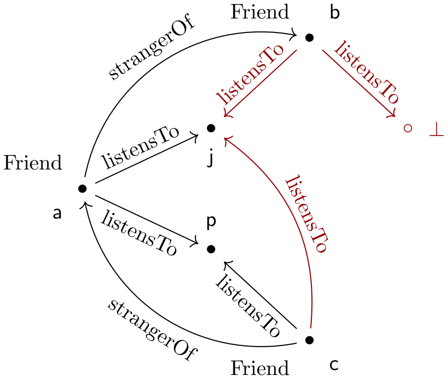


$$\forall x \; (\; \text{Friend}(x) \; \rightarrow \; \exists y \; \text{listensTo}(x, y) \;)$$

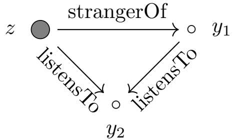
Model 1

$$\begin{array}{ccc} z & \mapsto & \textcolor{darkred}{\mathbf{a}} \\ y_1 & \mapsto & \mathsf{b} \\ y_2 & \mapsto & \mathsf{j} \end{array}$$

$$\begin{array}{ccc} z & \mapsto & \textcolor{darkred}{\mathbf{c}} \\ y_1 & \mapsto & \mathsf{a} \\ y_2 & \mapsto & \mathsf{j} \end{array}$$



Example of an OMQA setting



$$\forall x \; (\; \text{Friend}(x) \; \rightarrow \; \exists y \; \text{listensTo}(x, y) \;)$$

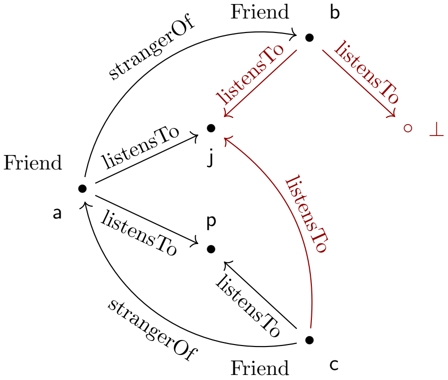
Model 1

$z \mapsto$ **a**
$y_1 \mapsto$ b
$y_2 \mapsto$ j

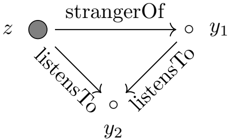
$z \mapsto$ **c**
$y_1 \mapsto$ a
$y_2 \mapsto$ j

$z \mapsto$ **c**
$y_1 \mapsto$ a
$y_2 \mapsto$ p

Answer: 2



Example of an OMQA setting



$$\forall x \; (\; \text{Friend}(x) \;\rightarrow\; \exists y \; \text{listensTo}(x,y) \;)$$

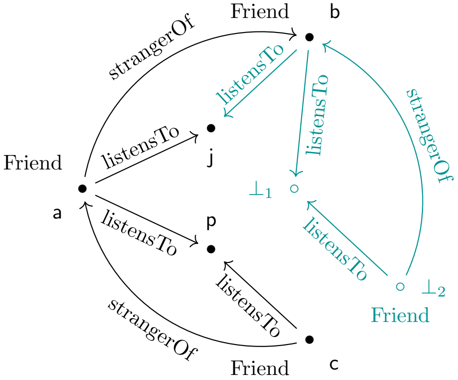
Model 2

$$\begin{array}{ccc} z & \mapsto & \mathbf{a} \\ y_1 & \mapsto & \mathsf{b} \\ y_2 & \mapsto & \mathsf{j} \end{array}$$

$$\begin{array}{ccc} z & \mapsto & \boldsymbol{\bot_2} \\ y_1 & \mapsto & \mathsf{b} \\ y_2 & \mapsto & \bot_1 \end{array}$$

$$\begin{array}{ccc} z & \mapsto & \mathbf{c} \\ y_1 & \mapsto & \mathsf{a} \\ y_2 & \mapsto & \mathsf{p} \end{array}$$

Answer: 3



Certain answers to *counting* CQs

Problem

How to reconcile these different answers?

Certain answers

Certain answer: interval $[m, M]$ such that in every model $\mathcal{I}$: $q^{\mathcal{I}} \in [m, M]$

Example (Continued)

Certain answer: $[1, +\infty]$
***Not** a certain answer:* $[3, +\infty]$
What about $[2, +\infty]$*?*

Related work

Kostylev, Reutter, JWS 2015	*Count* and *Cntd* CQs
Bienvenu et al., IJCAI 2020	Generalize both *Count* and *Cntd* CQs
Calvanese et al., IJCAI 2020	Restrictions for *Count* CQs Query rewriting techniques
Nikolaou et al., AIJ 2019	CQs over bag semantics
Calvanese et al., ONISW 2008	*Epistemic* aggregate queries
Feier et al., KR 2021	Counting the number of certain answers

DL-Lite

Focus on two well-known DL-Lite dialects

Concepts, Roles

Atomic roles	$P_1, P_2, \ldots$	Atomic concepts	$A_1, A_2, \ldots$
Positive roles	$R ::= P_j \mid P_j^-$	Positive concepts	$B ::= A_i \mid \exists R$
Roles	$E ::= R \mid \neg R$	Concepts	$C ::= B \mid \neg B$

with P_j^- being the inverse role and $\exists R$ the projection on the first variable

DL-Lite TBoxes

DL-Lite$_{\mathsf{core}}$: only concept inclusions of the form $B \sqsubseteq C$
DL-Lite$_{\mathcal{R}}$: also role inclusions of the form $R \sqsubseteq E$

Complexity of query answering

Combined complexity

Input Query q, TBox $\mathcal{T}$, ABox $\mathcal{A}$, and binary integer m.
Output Is $[m, +\infty]$ a certain answer for q and $(\mathcal{T}, \mathcal{A})$?

Data complexity ($\mathcal{T}$ and q are fixed)

Input ABox $\mathcal{A}$, and binary integer m.
Output Is $[m, +\infty]$ a certain answer for q and $(\mathcal{T}, \mathcal{A})$?

Complexity of query answering

Combined complexity

Input Query q, TBox $\mathcal{T}$, ABox $\mathcal{A}$, and binary integer m.
Output Is $[m, +\infty]$ a certain answer for q and $(\mathcal{T}, \mathcal{A})$?

Data complexity ($\mathcal{T}$ and q are fixed)

Input ABox $\mathcal{A}$, and binary integer m.
Output Is $[m, +\infty]$ a certain answer for q and $(\mathcal{T}, \mathcal{A})$?

Why $[m, +\infty]$ instead of $[m, M]$?

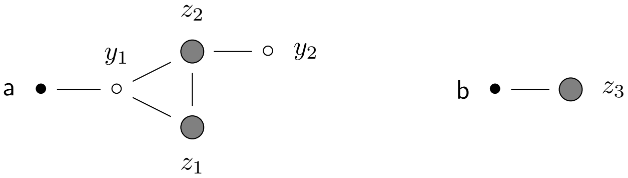
DL-Lite$_{\mathsf{core}}$, DL-Lite$_{\mathcal{R}}$: cannot restrict the size of models $\implies$ $M \in \{0, 1, +\infty\}$

Variant output: Is $[m, +\infty]$ *the tightest* certain answer for q and $(\mathcal{T}, \mathcal{A})$?

Complexity results

CCQs (generalize results from Kostylev and Reutter)
Rooted CCQs

	Data complexity	Combined complexity
DL-Lite$_{\mathsf{core}}$	coNP-complete	Π_2^p-hard PP-hard / coNEXP
DL-Lite$_{\mathcal{R}}$	coNP-complete	coNEXP-hard / coN2EXP



Rooted CCQ: each connected component contains a constant

Complexity results

Exhaustive rooted CCQ: rooted + no existential variables

Exhaustive rooted CCQs

	Data complexity	Combined complexity
DL-Lite$_{\mathsf{core}}$	TC^0-complete	PP-complete
DL-Lite$_{\mathcal{R}}$	coNP-complete	Π_2^p-hard PP-hard / coNEXP

Reminders: $\mathsf{TC}^0 \subseteq \mathsf{L} \subseteq \mathsf{P}$ $\mathsf{NP} \subseteq \mathsf{PP} \subseteq \mathsf{PSPACE}$

Complexity results

CCQs (generalize results from Kostylev and Reutter)
Rooted CCQs

	Data complexity	Combined complexity
DL-Lite$_{\mathsf{core}}$	coNP-complete	Π_2^p-hard, PP-hard / coNEXP
DL-Lite$_{\mathcal{R}}$	coNP-complete	coNEXP-hard / coN2EXP

Exhaustive rooted CCQs

	Data complexity	Combined complexity
DL-Lite$_{\mathsf{core}}$	TC^0-complete	PP-complete
DL-Lite$_{\mathcal{R}}$	coNP-complete	Π_2^p-hard, PP-hard / coNEXP

Reminders: $\mathsf{TC}^0 \subseteq \mathsf{L} \subseteq \mathsf{P}$ $\mathsf{NP} \subseteq \mathsf{PP} \subseteq \mathsf{PSPACE}$

Upper bounds: main ideas

Goal

Decide if $[m, +\infty]$ is a certain answer.

Idea: Look for a counter-model.

General setting	Bound the size of a possible counter-model (generalization of Kostylev and Reutter). $\rightsquigarrow$ Data complexity: in coNP $\rightsquigarrow$ Combined complexity: in $\mathsf{coN(2)EXP}$
Exhaustive rooted and DL-Lite$_{\mathsf{core}}$	The canonical model provides *the smallest* answer. $\rightsquigarrow$ Data complexity: in TC^0 $\rightsquigarrow$ Combined complexity: in PP

coNP-hardness w.r.t data complexity

Under the general setting, find an OMQ $(\mathcal{T}, q)$ s.t the following is coNP-hard (Kostylev Reutter).

Data complexity ($\mathcal{T}$ and q are fixed)

Input	ABox $\mathcal{A}$, and binary integer m.
Output	Is $[m, +\infty]$ a certain answer for q and $(\mathcal{T}, \mathcal{A})$?

coNP-hardness w.r.t data complexity

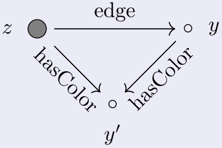
Replace: Friend by Vertex strangerOf by edge listensTo by hasColor

Example ontology becomes:

$$\forall x \; (\; \text{Vertex}(x) \;\rightarrow\; \exists y \; \text{hasColor}(x,y) \;)$$

Example query becomes:

$$\exists y \; \exists y' \; \exists z \qquad \text{edge}(z,y) \;\wedge\; \text{hasColor}(z,y') \;\wedge\; \text{hasColor}(y,y')$$



How many vertices z are connected with a vertex y of the same color y'?

coNP-hardness w.r.t data complexity

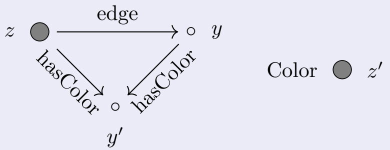
Replace: Friend by Vertex strangerOf by edge listensTo by hasColor

Example ontology extended:

$$\forall x \ (\ \text{Vertex}(x) \ \rightarrow \ \exists y \ \text{hasColor}(x, y) \ )$$
$$\forall x \ (\ \exists y \ \text{hasColor}(y, x) \ \rightarrow \ \text{Color}(x) \ )$$

Example query extended:

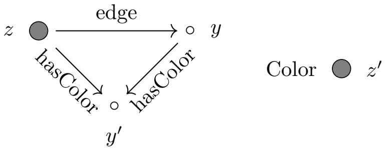
$$\exists y \ \exists y' \ \exists z \ \exists z' \ \text{edge}(z, y) \ \wedge \ \text{hasColor}(z, y') \ \wedge \ \text{hasColor}(y, y') \ \wedge \ \text{Color}(z')$$



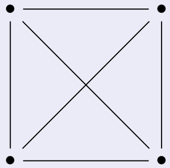
How many vertices z are connected with a vertex y of the same color y'
and how many colors z' are there?

Reduction from 3COL

$$\forall x \; (\; \text{Vertex}(x) \; \rightarrow \; \exists y \; \text{hasColor}(x,y) \;)$$
$$\forall x \; (\; \exists y \; \text{hasColor}(y,x) \; \rightarrow \; \text{Color}(x) \;)$$

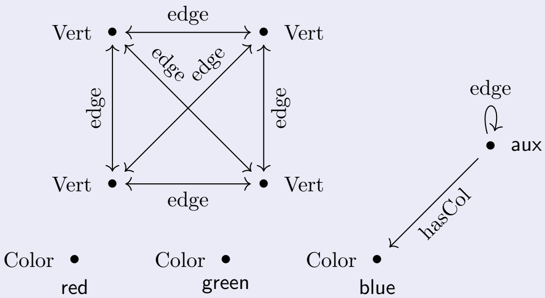


Input graph $\mathcal{G}$



$\rightsquigarrow$

ABox $\mathcal{A}_\mathcal{G}$



$$\boxed{[4, +\infty] \text{ is a certain answer for } q \text{ over } (\mathcal{T}, \mathcal{A}_\mathcal{G}) \iff \mathcal{G} \notin \text{3COL}}$$

Variant: Tightest certain answer

$[m, +\infty]$ is the tightest certain answer $\iff$ $[m, +\infty]$ is a certain answer and $[m+1, +\infty]$ is not

CCQs, rooted CCQs

	Original Data complexity		Variant Data complexity
DL-Lite$_{\mathsf{core}}$	coNP-complete	$\rightsquigarrow$	DP-complete
DL-Lite$_{\mathcal{R}}$	coNP-complete	$\rightsquigarrow$	DP-complete

Exhaustive rooted CCQs

	Original Data complexity		Variant Data complexity
DL-Lite$_{\mathsf{core}}$	TC^0-complete	$\rightsquigarrow$	TC^0-complete
DL-Lite$_{\mathcal{R}}$	coNP-complete	$\rightsquigarrow$	DP-complete

Cardinality Queries

Large gaps in our understanding of counting queries. Let's focus on a simple setting: Boolean atomic counting query!

Cardinality query

Concept cardinality query: $q_{\mathrm{C}} := \exists z\ \mathrm{C}(z)$.
Role cardinality query: $q_{\mathrm{S}} := \exists z_1\ \exists z_2\ \mathrm{S}(z_1, z_2)$.

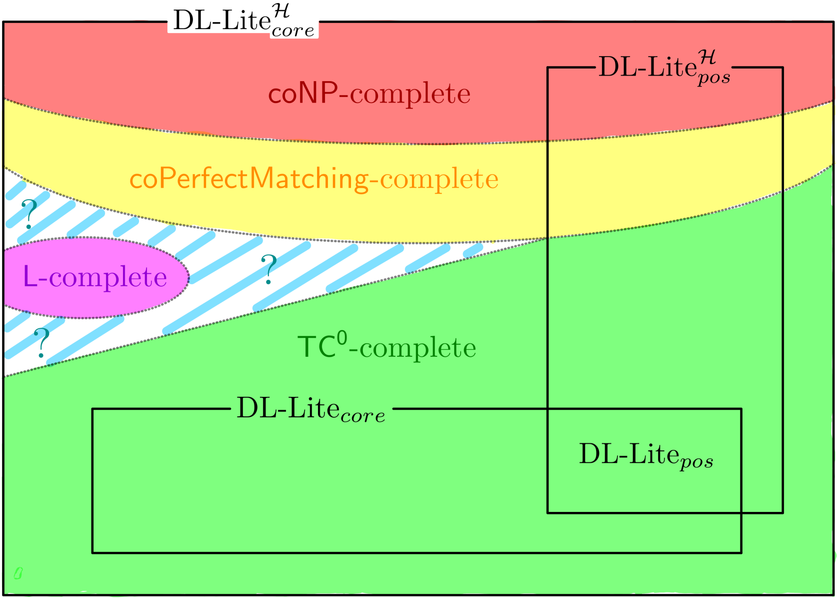
Complexity classification

	Concept Cardinality	Role Cardinality
DL-Lite$^{\mathcal{H}}_{core}$	TC^0-c | L-c | coNP-c | ?	TC^0-c | L-c | co-PM-c | coNP-c | ?
DL-Lite$^{\mathcal{H}}_{pos}$	TC^0-c	TC^0-c | co-PM-c | coNP-c
DL-Lite$_{core}$	TC^0-c	TC^0-c
DL-Lite$_{pos}$	TC^0-c	TC^0-c

Data complexity of cardinality queries for various DL-Lite dialects.

From now on: focus on the **Role** Cardinality Query $q_{\mathrm{S}} := \exists z_1 \; \exists z_2 \; \mathrm{S}(z_1, z_2)$.

Plan: Focus on Role Cardinality Query



Warm-up: DL-Lite$_{pos}$

Reminder: DL-Lite$_{pos}$ only allows concept inclusions.

Example

$$
\begin{array}{rclcl}
\mathrm{A} & \sqsubseteq & \exists \mathrm{S}^- & \quad & \forall x\ (\mathrm{A}(x) \rightarrow \exists y\ \mathrm{S}(y,x)) \\
\exists \mathrm{R}^- & \sqsubseteq & \exists \mathrm{R} & & \forall x\ (\exists y\ \mathrm{R}(y,x) \rightarrow \exists z\ \mathrm{R}(x,z)) \\
\exists \mathrm{R} & \sqsubseteq & \exists \mathrm{S} & & \forall x\ (\exists y\ \mathrm{R}(x,y) \rightarrow \exists z\ \mathrm{S}(x,z))
\end{array}
$$

$\mathrm{A}(\mathsf{a}_1)$ $\mathrm{A}(\mathsf{a}_2)$ $\mathrm{A}(\mathsf{a}_3)$ $\mathrm{R}(\mathsf{b},\mathsf{c})$ $\mathrm{S}(\mathsf{a}_1,\mathsf{b})$ $\mathrm{S}(\mathsf{a}_3,\mathsf{c})$

Goal

How to minimize the amount of matches?

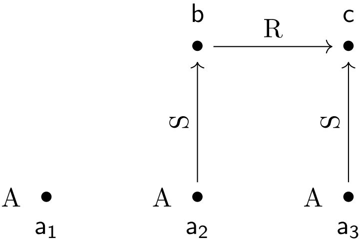
$\rightsquigarrow$ First idea: have a look at the canonical model.

Warm-up: DL-Lite$_{pos}$

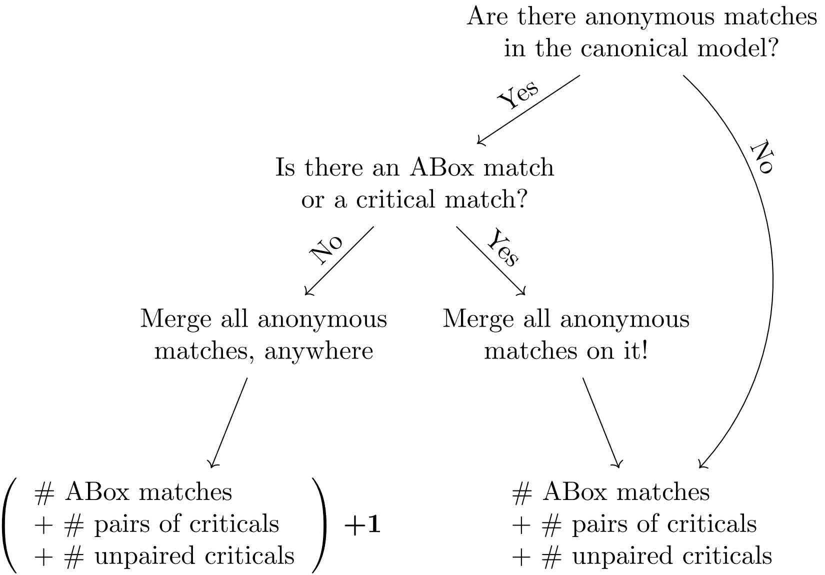
$$A \sqsubseteq \exists S^- \quad \exists R^- \sqsubseteq \exists R \quad \exists R \sqsubseteq \exists S$$
$$A(\mathsf{a}_1) \quad A(\mathsf{a}_2) \quad A(\mathsf{a}_3)$$
$$R(\mathsf{b}, \mathsf{c}) \quad S(\mathsf{a}_2, \mathsf{b}) \quad S(\mathsf{a}_3, \mathsf{c})$$

3 shapes of matches in $\mathcal{C}_\mathcal{K}$

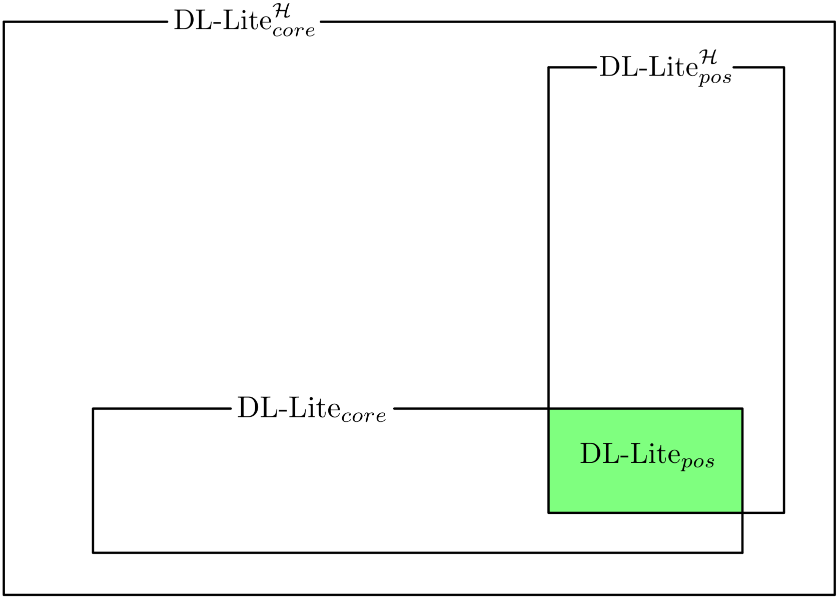
- ABox match: both variables map on individuals
- Anonymous match: no variable maps on individuals
- Critical match: one variable maps on an individual



DL-Lite$_{pos}$: Overview



Plan: warm-up is over!



Digression: canonical model for DL-Lite$^{\mathcal{H}}_{core}$ KBs

Reminder

There is a canonical model $\mathcal{C}_{\mathcal{K}}$: a model embedding in every model of $\mathcal{K}$.

$\rightsquigarrow$ Deciding if 1 is a certain answer is easy.

Digression: canonical model for DL-Lite$^{\mathcal{H}}_{core}$ KBs

Reminder

There is a canonical model $\mathcal{C}_{\mathcal{K}}$: a model embedding in every model of $\mathcal{K}$.

$\rightsquigarrow$ Deciding if 1 is a certain answer is easy.

Fact 1

Given $f : \mathcal{C}_{\mathcal{K}} \to \mathcal{I}$ an embedding, the induced interpretation $f(\mathcal{C}_{\mathcal{K}})$ is a model.

$\rightsquigarrow$ It is *sufficient* to understand "foldings" $f(\mathcal{C}_{\mathcal{K}})$ of the canonical model.

Digression: canonical model for DL-Lite$^{\mathcal{H}}_{core}$ KBs

Reminder

There is a canonical model $\mathcal{C}_{\mathcal{K}}$: a model embedding in every model of $\mathcal{K}$.

$\rightsquigarrow$ Deciding if 1 is a certain answer is easy.

Fact 1

Given $f : \mathcal{C}_{\mathcal{K}} \to \mathcal{I}$ an embedding, the induced interpretation $f(\mathcal{C}_{\mathcal{K}})$ is a model.

$\rightsquigarrow$ It is *sufficient* to understand "foldings" $f(\mathcal{C}_{\mathcal{K}})$ of the canonical model.

Digression in the digression: Fact 1 is non-trivial

Consider the $\mathcal{EL}$ knowledge base:

$$\mathrm{A} \sqsubseteq \exists \mathrm{R.B} \qquad \mathrm{A} \sqsubseteq \exists \mathrm{R.C} \qquad \mathrm{B} \sqcap \mathrm{C} \sqsubseteq \mathrm{D} \qquad\qquad \mathrm{A}(\mathsf{a})$$

Digression: canonical model for DL-Lite$^{\mathcal{H}}_{core}$ KBs

Reminder

There is a canonical model $\mathcal{C}_\mathcal{K}$: a model embedding in every model of $\mathcal{K}$.

$\rightsquigarrow$ Deciding if 1 is a certain answer is easy.

Fact 1

Given $f : \mathcal{C}_\mathcal{K} \to \mathcal{I}$ an embedding, the induced interpretation $f(\mathcal{C}_\mathcal{K})$ is a model.

$\rightsquigarrow$ It is *sufficient* to understand "foldings" $f(\mathcal{C}_\mathcal{K})$ of the canonical model.

Fact 2

Every match of q_S in $f(\mathcal{C}_\mathcal{K})$ is the image of a match of q_S in $\mathcal{C}_\mathcal{K}$.

$\rightsquigarrow$ It is *sufficient* to understand how to merge matches from $\mathcal{C}_\mathcal{K}$.

DL-Lite$_{core}$

Reminder: DL-Lite$_{core}$ allows concept inclusions and concept disjointness.

Example

$$\exists R^- \sqsubseteq \exists R \qquad\qquad \exists R \sqsubseteq \exists S$$

DL-Lite$_{core}$

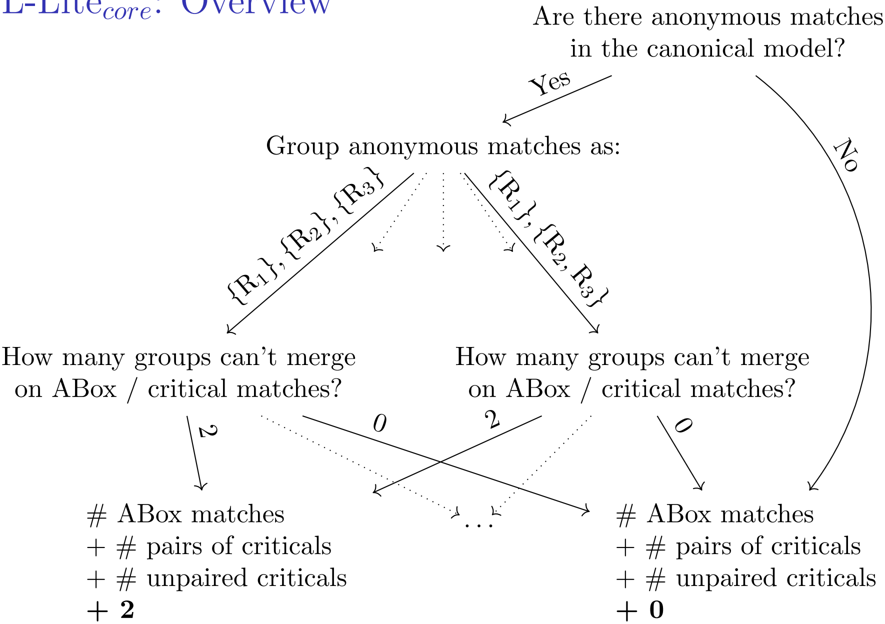
Reminder: DL-Lite$_{core}$ allows concept inclusions and concept disjointness.

Example

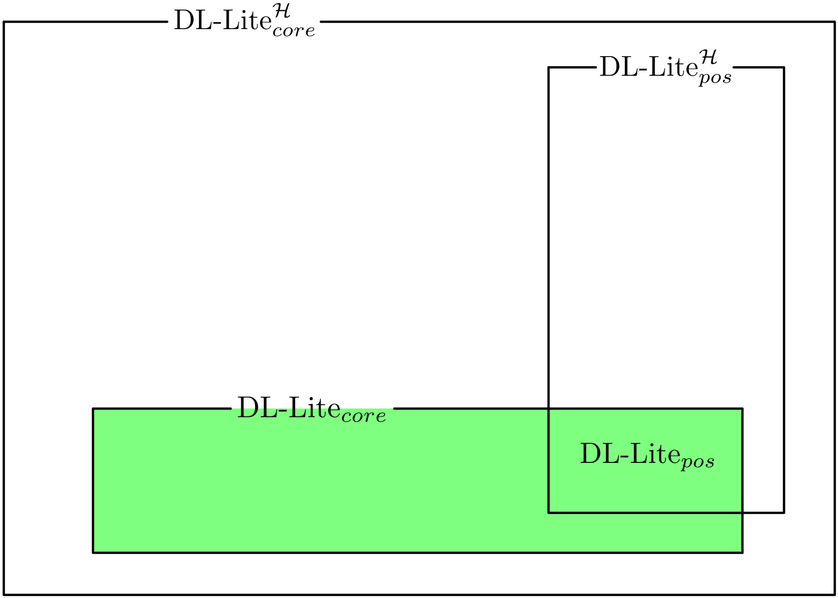
$$\begin{array}{rcl@{\qquad\qquad}rcl}
\exists R_1^- & \sqsubseteq & \exists R_1 & \exists R_1 & \sqsubseteq & \exists S \\
\exists R_2^- & \sqsubseteq & \exists R_2 & \exists R_2 & \sqsubseteq & \exists S \\
\exists R_3^- & \sqsubseteq & \exists R_3 & \exists R_3 & \sqsubseteq & \exists S \\
\exists R_1 & \sqsubseteq & \neg\exists R_2 & \exists R_3 & \sqsubseteq & \neg B
\end{array}$$

- ⇝ Anonymous matches from $\exists R_1$ and $\exists R_2$ cannot be merged.
- ⇝ Should those from $\exists R_3$ be merged with those from $\exists R_1$? from $\exists R_2$?
- ⇝ Not all ABox / critical matches make good merging target.

DL-Lite$_{core}$: Overview



Plan



DL-Lite$^{\mathcal{H}}_{core}$
DL-Lite$^{\mathcal{H}}_{pos}$
DL-Lite$_{core}$
DL-Lite$_{pos}$

DL-Lite$^{\mathcal{H}}_{pos}$: a coNP-hard situation

Reminder: DL-Lite$^{\mathcal{H}}_{pos}$ allows concept and role inclusions.

Never trust your subroles...

Answering q_{S} over the ontology $\{\ \exists \mathrm{R}^- \sqsubseteq \exists \mathrm{R}, \quad \mathrm{R} \sqsubseteq \mathrm{S}\ \}$ is coNP-complete.

$\rightsquigarrow$ Reduction from:

Set Cover (NP-complete)

Input: set $\mathcal{U}$, set of subsets $\mathcal{S} \subseteq 2^{\mathcal{U}}$, integer k.
Output: Is there a k-cover $\mathcal{C}$, i.e. a subset $\mathcal{C} \subseteq \mathcal{S}$ with $|\mathcal{C}| \leq k$ and $\bigcup \mathcal{C} = \mathcal{U}$.

DL-Lite$^{\mathcal{H}}_{pos}$: reduction from Set Cover

One instance of Set Cover

$\mathcal{U} = \{1, 2, 3, 4, 5\}$ $\quad$ $\mathcal{S} = \{\ \{1, 2\},\ \{3, 4\},\ \{4, 5\},\ \{1, 2, 3\}\ \}$ $\quad$ k

$(\mathcal{U}, \mathcal{S}, k) \in$ Set Cover

$\Longleftrightarrow$

$\Sigma_{s\in\mathcal{S}}|s| + k + 1$ is *not* a certain answer for q_{S} over $(\{\exists \mathrm{R}^- \sqsubseteq \exists \mathrm{R}, \mathrm{R} \sqsubseteq \mathrm{S}\}, \mathcal{A})$

DL-Lite$^{\mathcal{H}}_{pos}$: coNP-complete cases

Theorem (Sufficient condition for coNP-hardness)

Answering q_{S} over $\mathcal{T}$ is coNP*-complete if $\mathcal{T}$ admits a non-trivial propagation of either* S *or* S^{-}.

DL-Lite$^{\mathcal{H}}_{pos}$: coNP-complete cases

Theorem (Sufficient condition for coNP-hardness)

Answering q_{S} over $\mathcal{T}$ is coNP*-complete if $\mathcal{T}$ admits a non-trivial propagation of either* S *or* S^-.

Definition (Propagation of S)

A positive concept B *and positive roles* $\mathrm{R}_1, \mathrm{R}_2$ *such that $\mathcal{T}$ entails:*

$$\mathrm{B} \sqsubseteq \exists \mathrm{R}_1 \qquad \mathrm{R}_1 \sqsubseteq \mathrm{S} \qquad \exists \mathrm{R}_1^- \sqsubseteq \exists \mathrm{R}_2 \qquad \mathrm{R}_2 \sqsubseteq \mathrm{S}$$

DL-Lite$^{\mathcal{H}}_{pos}$: coNP-complete cases

Theorem (Sufficient condition for coNP-hardness)

Answering q_S over $\mathcal{T}$ is coNP*-complete if $\mathcal{T}$ admits a non-trivial propagation of either* S *or* S^-*.*

Definition (Propagation of S)

A positive concept B *and positive roles* $\mathrm{R}_1, \mathrm{R}_2$ *such that $\mathcal{T}$ entails:*

$$\mathrm{B} \sqsubseteq \exists \mathrm{R}_1 \qquad \mathrm{R}_1 \sqsubseteq \mathrm{S} \qquad \exists \mathrm{R}_1^- \sqsubseteq \exists \mathrm{R}_2 \qquad \mathrm{R}_2 \sqsubseteq \mathrm{S}$$

A role U makes the propagation of S trivial if one of the following holds:

- $\mathcal{T} \models \{\mathrm{B} \sqsubseteq \exists \mathrm{U}, \mathrm{U} \sqsubseteq \mathrm{S}, \mathrm{U} \sqsubseteq \mathrm{S}^-\}$;
- $\mathcal{T} \models \{\exists \mathrm{S}^- \sqsubseteq \exists \mathrm{U}, \mathrm{U} \sqsubseteq \mathrm{S}\}$ and either $\mathcal{T} \models \mathrm{U} \sqsubseteq \mathrm{S}^-$ or $\mathcal{T} \not\models \mathrm{R}_2 \sqsubseteq \mathrm{S}^-$;
- if $\mathrm{B} = \exists \mathrm{T}$ and $\mathrm{T} \sqsubseteq \mathrm{S}$, then $\mathcal{T} \models \{\exists \mathrm{T}^- \sqsubseteq \exists \mathrm{U}, \mathrm{U} \sqsubseteq \mathrm{S}\}$ and either $\mathcal{T} \models \mathrm{U} \sqsubseteq \mathrm{S}^-$ or $\mathcal{T} \not\models \mathrm{R}_2 \sqsubseteq \mathrm{S}^-$.

DL-Lite$^{\mathcal{H}}_{pos}$: RIP "pairing"

Double-agent...

Answering q_{S} over the ontology $\{ \quad \mathrm{B} \sqsubseteq \exists\mathrm{R}, \quad \mathrm{R} \sqsubseteq \mathrm{S}, \quad \mathrm{R} \sqsubseteq \mathrm{S}^{-} \quad \}$ is as hard as the Maximum Matching problem.

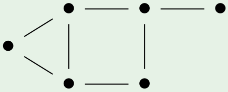
$\rightsquigarrow$ Reduction from:

Maximum Matching

Input: non-oriented graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, integer k.
Output: Is there a k-matching of $\mathcal{G}$, i.e. a subset $\mathcal{M} \subseteq \mathcal{E}$ of pairwise vertex-disjoint edges and with $|\mathcal{M}| \geq k$.

DL-Lite$^{\mathcal{H}}_{pos}$: reduction from Maximum Matching

A graph



$$(\mathcal{G}, k) \in \textsc{Maximum Matching}$$
$$\Longleftrightarrow$$
$|\mathcal{E}| - k + 1$ is *not* a certain answer for q_{S} over $(\{\mathrm{B} \sqsubseteq \exists\mathrm{R}, \mathrm{R} \sqsubseteq \mathrm{S}, \mathrm{R} \sqsubseteq \mathrm{S}^-\}, \mathcal{A})$

DL-Lite$^{\mathcal{H}}_{pos}$: Trichotomy

Definition (Non-trivial pairing)

Positive concept B *and positive role* R *such that:*

$$\mathcal{T} \models \mathrm{B} \sqsubseteq \exists \mathrm{R} \qquad \mathcal{T} \models \mathrm{R} \sqsubseteq \mathrm{S} \qquad \mathcal{T} \models \mathrm{R} \sqsubseteq \mathrm{S}^- \qquad \mathcal{T} \not\models \mathrm{S} \sqsubseteq \mathrm{S}^-$$

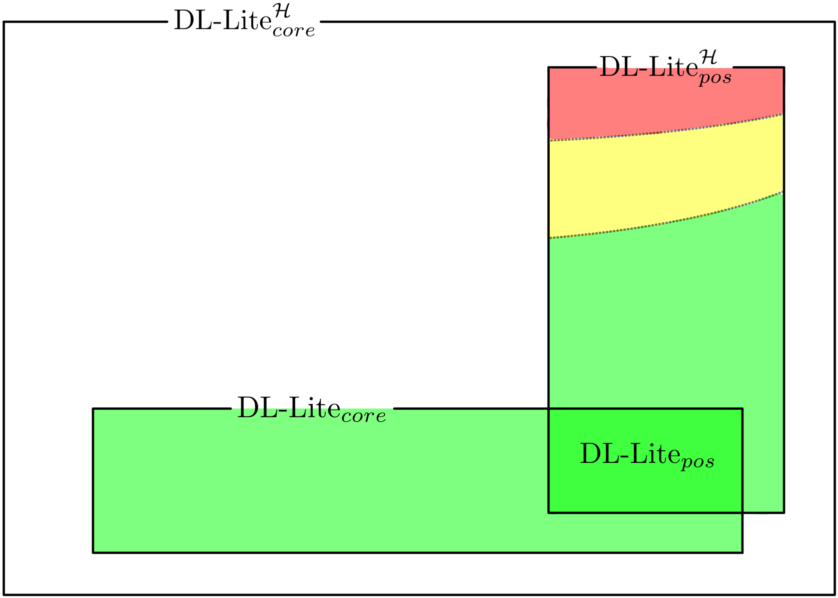
and if $\mathrm{B} = \exists \mathrm{T}$, *then either* $\mathcal{T} \not\models \mathrm{T} \sqsubseteq \mathrm{S}$ *or* $\mathcal{T} \not\models \mathrm{T} \sqsubseteq \mathrm{S}^-$.

Theorem (Trichotomy)

Answering q_{S} *over a DL-Lite*$^{\mathcal{H}}_{pos}$ *ontology* $\mathcal{T}$ *is:*

- coNP-*complete if* $\mathcal{T}$ *admits a non-trivial propagation of either* S *or* S^-;
- L-*equivalent to the complement of* PERFECT MATCHING *if it does not admit such a non-trivial propagation but admits a non-trivial pairing;*
- *in* TC^0 *otherwise.*

Plan: one last step...



DL-Lite$^{\mathcal{H}}_{core}$: a (coNP-)complete nightmare

Reminder: DL-Lite$^{\mathcal{H}}_{core}$ allows concept and role inclusions and disjointness.

- New sources of coNP hardness:

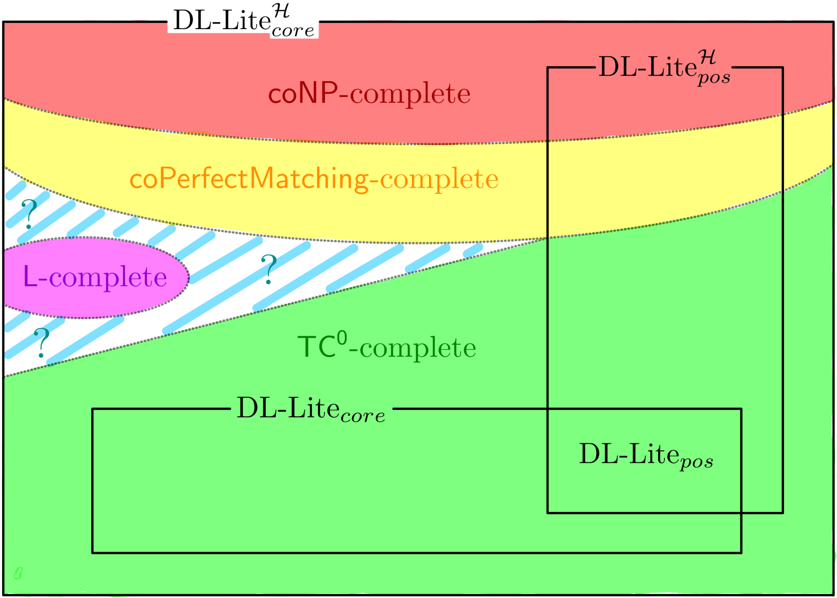
Answering q_{S} over $\{\ \mathrm{B} \sqsubseteq \exists \mathrm{U},\ \mathrm{U} \sqsubseteq \mathrm{S},\ \mathrm{C} \sqsubseteq \exists \mathrm{V},\ \mathrm{V} \sqsubseteq \mathrm{S},\ \exists \mathrm{U}^- \sqsubseteq \neg \exists \mathrm{V}^-\ \}$ is coNP-complete.

- Disjointness axioms: need to handle unsatisfiability issues

- New complexity case:

Answering q_{S} over $\{\ \mathrm{B} \sqsubseteq \exists \mathrm{R},\ \mathrm{R} \sqsubseteq \mathrm{S},\ \mathrm{R} \sqsubseteq \neg \mathrm{R}^-\ \}$ is L-complete.

Plan: we made it!



Ongoing & future work

Improve the complexity classification for DL-Lite$^{\mathcal{H}}_{core}$

- More than the 4 identified complexity cases?
- coNP-complete / in P dichotomy: adapt the propagation criteria?

Investigate other description logics / other queries

- Cardinality restrictions, functionality axioms: DL-Lite$_{\mathcal{F}}$, ...
- coNP-complete situations with $\mathcal{EL}$ ontology
- Classification for small queries, *e.g.* involving less than 3 variables
- Linear queries

Algorithms for tractable cases

- Investigate query rewriting techniques

Just-in-case definitions I

Canonical Model

The domain of $\mathcal{C}_\mathcal{K}$ contains $\mathsf{Ind}(\mathcal{A})$ and all words $\mathsf{a}\mathrm{R}_1 \dots \mathrm{R}_n$, with $\mathsf{a} \in \mathsf{Ind}(\mathcal{A})$, $\mathrm{R}_i \in \mathsf{N}_\mathsf{R}^\pm$, and $n \geqslant 1$, such that:

- $\mathcal{K} \models \exists \mathrm{R}_1(\mathsf{a})$ and there is no $\mathrm{R}_1(\mathsf{a}, \mathsf{b}) \in \mathcal{A}$;
- for $1 \leq i < n$, $\mathcal{T} \models \exists \mathrm{R}_i^- \sqsubseteq \exists \mathrm{R}_{i+1}$ and $\mathrm{R}_i^- \neq \mathrm{R}_{i+1}$.

Concept and role names are interpreted as follows:

$$\begin{aligned}
\mathrm{A}^{\mathcal{C}_\mathcal{K}} &= \{\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \mid \mathcal{K} \models \mathrm{A}(a)\} \\
&\quad \cup\ \{\mathsf{a}\mathrm{R}_1 \dots \mathrm{R}_n \in \Delta^{\mathcal{C}_\mathcal{K}} \setminus \mathsf{Ind}(\mathcal{A}) \mid \mathcal{T} \models \exists \mathrm{R}_n^- \sqsubseteq \mathrm{A}\} \\
\mathrm{P}^{\mathcal{C}_\mathcal{K}} &= \{(\mathsf{a}, \mathsf{b}) \mid \mathrm{P}(\mathsf{a}, \mathsf{b}) \in \mathcal{A}\} \\
&\quad \cup\ \{(e_1, e_2) \mid e_2 = e_1\mathrm{R} \text{ and } \mathcal{T} \models \mathrm{R} \sqsubseteq \mathrm{P}\} \\
&\quad \cup\ \{(e_2, e_1) \mid e_2 = e_1\mathrm{R} \text{ and } \mathcal{T} \models \mathrm{R} \sqsubseteq \mathrm{P}^-\}
\end{aligned}$$

Just-in-case definitions II

Type

A *type* for a TBox $\mathcal{T}$ is a subset of $\mathsf{sig}(\mathcal{T})^{\pm}_{\mathsf{C}}$. The set of all types is $\Theta_{\mathcal{T}} = 2^{\mathsf{sig}(\mathcal{T})^{\pm}_{\mathsf{C}}}$. We denote by $\theta_{\mathcal{K}}(d)$ the *type of a domain element d w.r.t. $\mathcal{K}$* and define it by: $\theta_{\mathcal{K}}(d) = \left\{ \mathrm{B} \in \mathsf{sig}(\mathcal{T})^{\pm}_{\mathsf{C}} \mid \mathcal{K} \models \mathrm{B}(d) \right\}$ if $d \in \mathsf{Ind}(\mathcal{A})$, else $\theta_{\mathcal{K}}(d) = \emptyset$.

	Undirected Forest Accessibility (UFA)
Input	Undirected acyclic graph $(\mathcal{V}, \mathcal{E})$ with 2 components, $s, t \in \mathcal{V}$
Output	Is t reachable from s ?