Containment for Conjunctive Queries with Negation

Highlights 2016

Gaetano Geck

— joint work with —

Bas Ketsman Frank Neven Thomas Schwentick

09.09.2016

technische universität dortmund

fakultät für informatik

What is Query Containment?

Problem: $\textsc{Containment}[\mathcal{C}]$
Input: Queries $Q_1, Q_2 \in \mathcal{C}$
Question: Is $Q_1(D) \subseteq Q_2(D)$ for every database D?

What is Query Containment?

Problem: $\textsc{Containment}[\mathcal{C}]$
Input: Queries $Q_1, Q_2 \in \mathcal{C}$
Question: Is $Q_1(D) \subseteq Q_2(D)$ for every database D?

Example 1.1

- Q_1: Is there a hotel that
  - is close to a cafe and
  - close to the university?

What is Query Containment?

Problem: $\textsc{Containment}[\mathcal{C}]$
Input: Queries $Q_1, Q_2 \in \mathcal{C}$
Question: Is $Q_1(D) \subseteq Q_2(D)$ for every database D?

Example 1.1

- Q_1: Is there a hotel that
  - is close to a cafe and
  - close to the university?
- Q_2:
  - Is there a hotel that is close to a cafe
  - and is there a hotel that is close to the university?

What is Query Containment?

Problem: $\textsc{Containment}[\mathcal{C}]$
Input: Queries $Q_1, Q_2 \in \mathcal{C}$
Question: Is $Q_1(D) \subseteq Q_2(D)$ for every database D?

Example 1.1

- $\boldsymbol{Q_1}$: Is there a hotel that
  - is close to a cafe and
  - close to the university?
- $\boldsymbol{Q_2}$:
  - Is there a hotel that is close to a cafe
  - and is there a hotel that is close to the university?

Clear: Q_1 is **contained** in Q_2 (Notation: $Q_1 \subseteq Q_2$)

Query Classes under Consideration

Conjunctive Query (CQ):

$$Q: \quad \texttt{Ans}(x, z) \quad \leftarrow \quad R(x, y), \quad R(y, z)$$

Query Classes under Consideration

Conjunctive Query (CQ):

$$Q: \quad \underbrace{\texttt{Ans}(x, z)}_{\boldsymbol{head}_Q} \quad \leftarrow \quad R(x, y), \quad R(y, z)$$

Query Classes under Consideration

Conjunctive Query (CQ):

$$Q: \ \texttt{Ans}(x,z) \ \leftarrow \ \underbrace{R(x,y),}_{\boldsymbol{pos}_Q} \ \underbrace{R(y,z)}_{\boldsymbol{pos}_Q}$$

Query Classes under Consideration

Conjunctive Query with Negation ($CQ^{\neg}$):

$$Q: \ \texttt{Ans}(x,z) \ \leftarrow \ R(x,y), \ \ R(y,z), \ \ \neg \underbrace{S(x,x,y)}_{\boldsymbol{neg}_Q}$$

Query Classes under Consideration

Conjunctive Query with Negation ($CQ^{\neg}$):

$$Q: \ \texttt{Ans}(x, z) \ \leftarrow \ R(x, y), \ \ R(y, z), \ \ \neg S(x, x, y)$$

Database

R

(1,	2)
(1,	3)
(2,	3)

S

(1,	1,	3)
(3,	4,	5)
(4,	5,	3)

Query Classes under Consideration

Conjunctive Query with Negation ($CQ^{\neg}$):

$$Q: \quad \texttt{Ans}(x,z) \quad \leftarrow \quad R(x,y), \quad R(y,z), \quad \neg S(x,x,y)$$

$$V(Q): \quad \texttt{Ans}(1,3) \qquad R(1,2) \quad R(2,3) \quad S(1,1,2)$$

Database

R

(1,	2)
(1,	3)
(2,	3)

S

(1,	1,	3)
(3,	4,	5)
(4,	5,	3)

Preview

Deciding $\textsc{Containment}[\mathcal{C}]$**:**

$\mathcal{C}$	**complexity**	**intuition**
$\mathbf{CQ}$	NP	fixed polynomial size canonical database
$\mathbf{CQ}^{\neg}$ bounded	Π_2^p	exponentially many polynomial size canonical databases
$\mathbf{CQ}^{\neg}$ general	coNEXPTIME	exponentially many exponential size canonical databases

Preview

Deciding $\overline{\textsc{Containment}}[\mathcal{C}]$:

$\mathcal{C}$	complexity	intuition
$\mathbf{CQ}$	coNP	fixed polynomial size counterexample
$\mathbf{CQ}^{\neg}$ bounded	Σ_2^p	some polynomial size counterexample
$\mathbf{CQ}^{\neg}$ general	NEXPTIME	some exponential size counterexample

Showing Non-Containment

$$Q_1 \nsubseteq Q_2$$

Showing Non-Containment

$$Q_1 \not\subseteq Q_2$$

$$\Leftrightarrow$$

$$\exists D \exists f$$

Showing Non-Containment

$$Q_1 \not\subseteq Q_2$$

$$\Leftrightarrow$$

$$\exists D \exists f \left(f \in Q_1(D) \qquad\qquad\qquad \right)$$

Showing Non-Containment

$$Q_1 \nsubseteq Q_2$$

$$\Leftrightarrow$$

$$\exists D \exists f \left(f \in Q_1(D) \land f \notin Q_2(D) \right)$$

$\overline{\text{Containment}}[\boldsymbol{CQ}]$

$$Q_1: \quad \texttt{Ans}(x,y) \quad \leftarrow \quad R(x,y), \quad R(y,z), \quad S(x,x,y)$$

$$Q_2: \quad \texttt{Ans}(u,w) \quad \leftarrow \quad R(u,v), \quad R(v',w), \quad S(u,u,v), \quad S(v',v',u)$$

Containment[CQ]

$$
\begin{array}{lccccc}
Q_1: & \texttt{Ans}(x,y) & \leftarrow & R(x,y), & R(y,z), & S(x,x,y) \\
 & \downarrow & & \downarrow & \downarrow & \downarrow \\
\mathbf{id}(Q_1): & \texttt{Ans}(x,y) & & R(x,y) & R(y,z) & S(x,x,y)
\end{array}
$$

$$
Q_2: \quad \texttt{Ans}(u,w) \quad \leftarrow \quad R(u,v), \quad R(v',w), \quad S(u,u,v), \quad S(v',v',u)
$$

Containment[CQ]

$$Q_1: \quad \texttt{Ans}(x,y) \leftarrow R(x,y),\ R(y,z),\ S(x,x,y)$$

$$\mathbf{id}(Q_1): \quad \texttt{Ans}(x,y) \quad R(x,y) \quad R(y,z) \quad S(x,x,y)$$

counterexample: canonical database D_{Q_1}

R	
(x,	y)
(y,	z)

S		
(x,	x,	y)

$$Q_2: \quad \texttt{Ans}(u,w) \leftarrow R(u,v),\ R(v',w),\ S(u,u,v),\ S(v',v',u)$$

Containment[CQ]

$$Q_1 : \ \texttt{Ans}(x,y) \leftarrow R(x,y),\ R(y,z),\ S(x,x,y)$$

$$\mathbf{id}(Q_1) : \ \texttt{Ans}(x,y) \quad R(x,y) \quad R(y,z) \quad S(x,x,y)$$

counterexample: canonical database D_{Q_1}

R	
(x,	y)
(y,	z)

S		
(x,	x,	y)

$$V_2(Q_2) : \ \texttt{Ans}(x,z) \quad R(x,y) \quad R(y,z) \quad S(x,x,y) \quad S(y,y,x)$$

$$Q_2 : \ \texttt{Ans}(u,w) \leftarrow R(u,v),\ R(v',w),\ S(u,u,v),\ S(v',v',u)$$

$\overline{\text{Containment}}[CQ]$: Complexity

Theorem 3.1 (Ramakrishnan et al., 1989)

Let Q_1 and Q_2 be conjunctive queries.

Then $Q_1 \nsubseteq Q_2$ if and only if $\boldsymbol{head}_{Q_1} \notin Q_2(\boldsymbol{D}_{Q_1})$.

$\overline{\textsc{Containment}}[\boldsymbol{CQ}]$: Complexity

Theorem 3.1 (Ramakrishnan et al., 1989)

Let Q_1 and Q_2 be conjunctive queries.
Then $Q_1 \nsubseteq Q_2$ if and only if $\boldsymbol{head}_{Q_1} \notin Q_2(\boldsymbol{D}_{Q_1})$.

Theorem 3.2 (Chandra, Merlin; 1977)

Deciding $\overline{\textsc{Containment}}[\boldsymbol{CQ}]$ is coNP*-complete.*

$\overline{\textsc{Containment}}[\boldsymbol{CQ}]$: Complexity

Theorem 3.1 (Ramakrishnan et al., 1989)

Let Q_1 and Q_2 be conjunctive queries.
Then $Q_1 \nsubseteq Q_2$ if and only if $\boldsymbol{head}_{Q_1} \notin Q_2(\boldsymbol{D}_{Q_1})$.

Theorem 3.2 (Chandra, Merlin; 1977)

Deciding $\overline{\textsc{Containment}}[\boldsymbol{CQ}]$ is coNP*-complete.*

Proof (idea)

- **Upper bound:** For every valuation V_2 for Q_2

$\overline{\textsc{Containment}}[\boldsymbol{CQ}]$: Complexity

Theorem 3.1 (Ramakrishnan et al., 1989)

Let Q_1 and Q_2 be conjunctive queries.

Then $Q_1 \nsubseteq Q_2$ if and only if $\boldsymbol{head}_{Q_1} \notin Q_2(\boldsymbol{D}_{Q_1})$.

Theorem 3.2 (Chandra, Merlin; 1977)

Deciding $\overline{\textsc{Containment}}[\boldsymbol{CQ}]$ is coNP*-complete.*

Proof (idea)

- **Upper bound:** For every valuation V_2 for Q_2 check that
  - $V_2(\boldsymbol{pos}_{Q_2}) \nsubseteq \boldsymbol{D}_{Q_2}$

$\overline{\text{Containment}}[\boldsymbol{CQ}]$: Complexity

Theorem 3.1 (Ramakrishnan et al., 1989)

Let Q_1 and Q_2 be conjunctive queries.

Then $Q_1 \nsubseteq Q_2$ if and only if $\boldsymbol{head}_{Q_1} \notin Q_2(\boldsymbol{D}_{Q_1})$.

Theorem 3.2 (Chandra, Merlin; 1977)

Deciding $\overline{\text{Containment}}[\boldsymbol{CQ}]$ is coNP*-complete.*

Proof (idea)

- **Upper bound:** For every valuation V_2 for Q_2 check that
  - $V_2(\boldsymbol{pos}_{Q_2}) \nsubseteq \boldsymbol{D}_{Q_2}$ or
  - $V_2(\boldsymbol{head}_{Q_2}) \neq \boldsymbol{head}_{Q_1}$

$\overline{\textsc{Containment}}[\boldsymbol{CQ}]$: Complexity

Theorem 3.1 (Ramakrishnan et al., 1989)

Let Q_1 and Q_2 be conjunctive queries.

Then $Q_1 \nsubseteq Q_2$ <u>if and only if</u> $\boldsymbol{head}_{Q_1} \notin Q_2(\boldsymbol{D}_{Q_1})$.

Theorem 3.2 (Chandra, Merlin; 1977)

Deciding $\overline{\textsc{Containment}}[\boldsymbol{CQ}]$ is coNP*-complete.*

Proof (idea)

- **Upper bound:** For every valuation V_2 for Q_2 check that
  - $V_2(\boldsymbol{pos}_{Q_2}) \nsubseteq \boldsymbol{D}_{Q_2}$ <u>or</u>
  - $V_2(\boldsymbol{head}_{Q_2}) \neq \boldsymbol{head}_{Q_1}$
- **Lower bound:** Reduction from $\overline{\textsc{Col}}$

$\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$

$$Q_1: \quad \texttt{Ans}() \quad \leftarrow \quad R(x_1, x_2, x_3)$$

$$Q_2: \quad \texttt{Ans}() \quad \leftarrow \quad R(y_1, y_2, y_3), \quad \neg R(y_2, y_3, y_1)$$

$\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$

$$Q_1: \quad \texttt{Ans}() \quad \leftarrow \quad R(x_1, x_2, x_3)$$

$$\mathbf{id}(Q_1): \quad \texttt{Ans}() \qquad\quad R(x_1, x_2, x_3)$$

$$Q_2: \quad \texttt{Ans}() \quad \leftarrow \quad R(y_1, y_2, y_3), \quad \neg R(y_2, y_3, y_1)$$

Containment[$CQ^{\neg}$]

$$Q_1 : \quad \texttt{Ans}() \quad \leftarrow \quad R(x_1, x_2, x_3)$$

$$\downarrow \qquad\qquad \downarrow$$

$$\mathbf{id}(Q_1) : \quad \texttt{Ans}() \qquad R(x_1, x_2, x_3)$$

possible counterexample

R		
$(x_1,$	$x_2,$	$x_3)$

$$Q_2 : \quad \texttt{Ans}() \quad \leftarrow \quad R(y_1, y_2, y_3), \quad \neg R(y_2, y_3, y_1)$$

CONTAINMENT[$\boldsymbol{CQ^{\neg}}$]

$$Q_1: \ \texttt{Ans}() \leftarrow R(x_1, x_2, x_3)$$

$$\downarrow \qquad \downarrow$$

$$\mathbf{id}(Q_1): \ \texttt{Ans}() \quad R(x_1, x_2, x_3)$$

possible counterexample

R		
$(x_1,$	$x_2,$	$x_3)$

$$V_2(Q_2): \ \texttt{Ans}() \quad R(x_1, x_2, x_3) \quad R(x_2, x_3, x_1)$$

$$\uparrow \qquad \uparrow \qquad \uparrow$$

$$Q_2: \ \texttt{Ans}() \leftarrow R(y_1, y_2, y_3), \ \neg R(y_2, y_3, y_1)$$

CONTAINMENT[$CQ^{\neg}$]

$$Q_1 : \texttt{Ans}() \leftarrow R(x_1, x_2, x_3)$$

$$\mathbf{id}(Q_1) : \texttt{Ans}() \quad R(x_1, x_2, x_3)$$

possible counterexample

R		
$(x_1,$	$x_2,$	$x_3)$
$(x_2,$	$x_3,$	$x_1)$

$R(x_2, x_3, x_1)$

$$Q_2 : \texttt{Ans}() \leftarrow R(y_1, y_2, y_3), \quad \neg R(y_2, y_3, y_1)$$

Containment[$CQ^{\neg}$]

$$Q_1: \quad \texttt{Ans}() \quad \leftarrow \quad R(x_1, x_2, x_3)$$

$$\downarrow \qquad\qquad \downarrow$$

$$\mathbf{id}(Q_1): \quad \texttt{Ans}() \qquad R(x_1, x_2, x_3)$$

possible counterexample

R		
$(x_1,$	$x_2,$	$x_3)$
$(x_2,$	$x_3,$	$x_1)$

$$V_2'(Q_2): \quad \texttt{Ans}() \qquad R(x_2, x_3, x_1) \quad R(x_3, x_1, x_2)$$

$$\uparrow \qquad\qquad \uparrow \qquad\qquad \uparrow$$

$$Q_2: \quad \texttt{Ans}() \quad \leftarrow \quad R(y_1, y_2, y_3), \quad \neg R(y_2, y_3, y_1)$$

Containment[$CQ^{\neg}$]

$$Q_1 : \texttt{Ans}() \leftarrow R(x_1, x_2, x_3)$$

$$\mathbf{id}(Q_1) : \texttt{Ans}() \quad R(x_1, x_2, x_3)$$

possible counterexample

R		
$(x_1,$	$x_2,$	$x_3)$
$(x_2,$	$x_3,$	$x_1)$
$(x_3,$	$x_1,$	$x_2)$

$R(x_3, x_1, x_2)$

$$Q_2 : \texttt{Ans}() \leftarrow R(y_1, y_2, y_3), \quad \neg R(y_2, y_3, y_1)$$

Containment[$CQ^{\neg}$]

$$Q_1 : \texttt{Ans}() \leftarrow R(x_1, x_2, x_3)$$

$$\downarrow \qquad \downarrow$$

$$\mathbf{id}(Q_1) : \texttt{Ans}() \quad R(x_1, x_2, x_3)$$

possible counterexample

R		
$(x_1,$	$x_2,$	$x_3)$
$(x_2,$	$x_3,$	$x_1)$
$(x_3,$	$x_1,$	$x_2)$

$$V_2''(Q_2) : \texttt{Ans}() \quad R(x_3, x_1, x_2) \quad R(x_1, x_2, x_3)$$

$$\uparrow \qquad \uparrow \qquad \uparrow$$

$$Q_2 : \texttt{Ans}() \leftarrow R(y_1, y_2, y_3), \quad \neg R(y_2, y_3, y_1)$$

CONTAINMENT[$CQ^{\neg}$]: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample

$\overline{\text{CONTAINMENT}}[\boldsymbol{CQ^{\neg}}]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered

CONTAINMENT[$\boldsymbol{CQ}^{\neg}$]: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $\boldsymbol{var}(Q_1)$, map equal variables to same value

$\overline{\text{Containment}}[\boldsymbol{CQ}^\neg]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $\boldsymbol{var}(Q_1)$, map equal variables to same value
  - exponential number of equivalence relations

$\overline{\text{CONTAINMENT}}[\boldsymbol{CQ}^\neg]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $\boldsymbol{var}(Q_1)$, map equal variables to same value
  - exponential number of equivalence relations

Theorem 4.1 (Levy, Sagiv; 1993 and Mugnier et al.; 2007)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^\neg$ over relations of bounded arity.

Deciding $\overline{\text{CONTAINMENT}}[\boldsymbol{CQ}^\neg]$ is Σ_2^p-complete.

$\overline{\text{Containment}}[\boldsymbol{CQ}^\neg]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $\boldsymbol{var}(Q_1)$, map equal variables to same value
  - exponential number of equivalence relations

Theorem 4.1 (Levy, Sagiv; 1993 and Mugnier et al.; 2007)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^\neg$ over relations of bounded arity.
Deciding $\overline{\text{Containment}}[\boldsymbol{CQ}^\neg]$ is Σ_2^p-complete.

Proof (idea)

- **Upper bound:**
  - Guess counterexample $\boldsymbol{D}$ of size $|\boldsymbol{D}| \le |\boldsymbol{var}(Q_1)|^{k_1+\cdots+k_n}$

$\overline{\text{Containment}}[\boldsymbol{CQ}^\neg]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $\boldsymbol{var}(Q_1)$, map equal variables to same value
  - exponential number of equivalence relations

Theorem 4.1 (Levy, Sagiv; 1993 and Mugnier et al.; 2007)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^\neg$ over relations of bounded arity.
Deciding $\overline{\text{Containment}}[\boldsymbol{CQ}^\neg]$ is Σ_2^p-complete.

Proof (idea)

- **Upper bound:**
  - Guess counterexample $\boldsymbol{D}$ of size $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{k_1+\dots+k_n}$ for relations $R_1, \dots, R_n$ of arities $k_1, \dots, k_n$,

$\overline{\text{Containment}}[\boldsymbol{CQ}^\neg]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $\boldsymbol{var}(Q_1)$, map equal variables to same value
  - exponential number of equivalence relations

Theorem 4.1 (Levy, Sagiv; 1993 and Mugnier et al.; 2007)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^\neg$ over relations of bounded arity.
Deciding $\overline{\text{Containment}}[\boldsymbol{CQ}^\neg]$ is Σ_2^p-complete.

Proof (idea)

- **Upper bound:**
  - Guess counterexample $\boldsymbol{D}$ of size $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{k_1+\dots+k_n}$ for relations $R_1, \dots, R_n$ of arities $k_1, \dots, k_n$,
  - guess satisfying valuation V_1 for Q_1

$\overline{\text{CONTAINMENT}}[\boldsymbol{CQ}^{\neg}]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $\boldsymbol{var}(Q_1)$, map equal variables to same value
  - exponential number of equivalence relations

Theorem 4.1 (Levy, Sagiv; 1993 and Mugnier et al.; 2007)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^{\neg}$ over relations of bounded arity.

Deciding $\overline{\text{CONTAINMENT}}[\boldsymbol{CQ}^{\neg}]$ is Σ_2^p-complete.

Proof (idea)

- **Upper bound:**
  - Guess counterexample $\boldsymbol{D}$ of size $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{k_1+\cdots+k_n}$ for relations $R_1, \ldots, R_n$ of arities $k_1, \ldots, k_n$,
  - guess satisfying valuation V_1 for Q_1
  - Check all valuations V_2 for Q_2 (as before)

$\overline{\text{Containment}}[CQ^\neg]$: Complexity (1/2)

Insight:

- Need to consider all extensions of counterexample
- Also: multiple counterexamples have to be considered
  - for every equivalence relation over $var(Q_1)$, map equal variables to same value
  - exponential number of equivalence relations

Theorem 4.1 (Levy, Sagiv; 1993 and Mugnier et al.; 2007)

Let $Q_1, Q_2 \in CQ^\neg$ over relations of bounded arity.

Deciding $\overline{\text{Containment}}[CQ^\neg]$ is Σ_2^p-complete.

Proof (idea)

- **Upper bound:**
  - Guess counterexample D of size $|D| \leq |var(Q_1)|^{k_1+\dots+k_n}$ for relations $R_1, \dots, R_n$ of arities $k_1, \dots, k_n$,
  - guess satisfying valuation V_1 for Q_1
  - Check all valuations V_2 for Q_2 (as before)
- **Lower bound:** reduction from GeneralisedRamseyNumber

$\overline{\text{Containment}}[CQ^{\neg}]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

$\overline{\text{Containment}}[CQ^{\neg}]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

- arities $k_1, \ldots, k_n$ of relations depend on $|Q_1|$

$\overline{\text{CONTAINMENT}}[\boldsymbol{CQ}^{\neg}]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

- arities $k_1, \ldots, k_n$ of relations depend on $|Q_1|$
- size of counterexample: $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{|Q_1|}$

$\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

- arities $k_1, \ldots, k_n$ of relations depend on $|Q_1|$
- size of counterexample: $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{|Q_1|}$

Theorem 4.2 (ICDT'16)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^{\neg}$.
Deciding $\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$ is NEXPTIME*-complete in general.*

$\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

- arities $k_1, \ldots, k_n$ of relations depend on $|Q_1|$
- size of counterexample: $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{|Q_1|}$

Theorem 4.2 (ICDT'16)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^{\neg}$.
Deciding $\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$ is NEXPTIME*-complete in general.*

Proof (idea)

- **Upper bound:** Algorithm as before

$\overline{\text{CONTAINMENT}}[\boldsymbol{CQ}^\neg]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

- arities $k_1, \ldots, k_n$ of relations depend on $|Q_1|$
- size of counterexample: $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{|Q_1|}$

Theorem 4.2 (ICDT'16)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^\neg$*.*
Deciding $\overline{\text{CONTAINMENT}}[\boldsymbol{CQ}^\neg]$ *is* NEXPTIME*-complete in general.*

Proof (idea)

- **Upper bound:** Algorithm as before
- **Lower bound:** holds even for *boolean* $\mathrm{CQ}^\neg$

$\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

- arities $k_1, \ldots, k_n$ of relations depend on $|Q_1|$
- size of counterexample: $|\boldsymbol{D}| \leq |\boldsymbol{var}(Q_1)|^{|Q_1|}$

Theorem 4.2 (ICDT'16)

Let $Q_1, Q_2 \in \boldsymbol{CQ}^{\neg}$.
Deciding $\overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}]$ is NEXPTIME*-complete in general.*

Proof (idea)

- **Upper bound:** Algorithm as before
- **Lower bound:** holds even for *boolean* $\mathrm{CQ}^{\neg}$

1) $\text{Succinct-3-Col} \leq_p \overline{\text{Containment}}[\boldsymbol{CQ}^{\neg}, \boldsymbol{UCQ}^{\neg}]$

$\overline{\text{Containment}}[\boldsymbol{CQ^\neg}]$: Complexity (2/2)

Question: What if schema not fixed, arities not bounded?

- arities $k_1, \ldots, k_n$ of relations depend on $|Q_1|$
- size of counterexample: $|\boldsymbol{D}| \leq \left|\boldsymbol{var}(Q_1)\right|^{|Q_1|}$

Theorem 4.2 (ICDT'16)

Let $Q_1, Q_2 \in \boldsymbol{CQ^\neg}$.
Deciding $\overline{\text{Containment}}[\boldsymbol{CQ^\neg}]$ is NEXPTIME*-complete in general.*

Proof (idea)

- **Upper bound:** Algorithm as before
- **Lower bound:** holds even for *boolean* $\mathrm{CQ}^\neg$

1) $\textsc{Succinct-3-Col} \leq_p \overline{\text{Containment}}[\boldsymbol{CQ^\neg}, \boldsymbol{UCQ^\neg}]$

2) $\overline{\text{Containment}}[\boldsymbol{CQ^\neg}, \boldsymbol{UCQ^\neg}] \leq_p \overline{\text{Containment}}[\boldsymbol{CQ^\neg}]$

Summary

Deciding $\textsc{Containment}[\mathcal{C}]$**:**

$\mathcal{C}$	complexity	intuition
CQ	NP	fixed polynomial size canonical database
$\mathbf{CQ}^\neg$ bounded	Π_2^p	exponentially many polynomial size canonical databases
$\mathbf{CQ}^\neg$ general	coNEXPTIME	exponentially many exponential size canonical databases

Merci! Dank u wel!

References

Gaetano Geck, Bas Ketsman, Frank Neven and Thomas Schwentick.
Parallel-Correctness and Containment
for Conjunctive Queries with Union and Negation.
ICDT, 2016.

Ashok K. Chandra, Philip M. Merlin.
Optimal Implementation of Conjunctive Queries in Relational Data Bases.
In STOC, pages 77 – 90, 1977.

Marie-Laure Mugnier, Geneviève Simonet, Michaël Thomazo
On the Complexity of Entailment in Existential Conjunctive
First-Order Logic with Atomic Negation.
In Inf. Comput. (IANDC), pages 8 – 31, 2012.

Alon Y. Levy, Yehoshua Sagiv.
Queries Independent of Updates.
VLDB, pages 171 – 181, 1993.