Classification of Annotation Semirings over Query Containment

Егор В. Костылев      Juan L. Reutter      András Z. Salamon

LFCS, University of Edinburgh
Relational Database annotation
## Relational Database annotation: Comments

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td></td>
<td>Top mark</td>
</tr>
<tr>
<td>Jane</td>
<td>Physics</td>
<td></td>
<td>Wants TA</td>
</tr>
<tr>
<td>Anne</td>
<td>History</td>
<td></td>
<td>Class Rep.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likes</th>
<th>Student</th>
<th>Course</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td></td>
<td>Wants TA</td>
</tr>
<tr>
<td>Anne</td>
<td>Literature</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Relational Database annotation: Comments

**Takes**

<table>
<thead>
<tr>
<th>Student</th>
<th>Course</th>
<th>Top mark</th>
<th>Wants TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jane</td>
<td>Physics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>History</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Likes**

<table>
<thead>
<tr>
<th>Student</th>
<th>Course</th>
<th>Wants TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>Literature</td>
<td></td>
</tr>
</tbody>
</table>

```
SELECT Student, Course
FROM Takes, Likes
WHERE Takes.S = Likes.S
AND Takes.C = Likes.C
```
Relational Database annotation: *Belief*

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>Teach. Office</th>
<th>Stud. Union</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jane</td>
<td>Physics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>History</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likes</th>
<th>Student</th>
<th>Course</th>
<th>Stud. Union</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td></td>
<td>Stud. Union</td>
</tr>
<tr>
<td>Anne</td>
<td>Literature</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


(Jane, Algebra): Stud. Union
Relational Database annotation: *Bag Semantics*

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Jane</td>
<td>Physics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>History</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likes</th>
<th>Student</th>
<th>Course</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>Literature</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

```
SELECT Student, Course
FROM Takes, Likes
WHERE Takes.S = Likes.S
AND Takes.C = Likes.C
```

*(Jane, Algebra)*:

\[2 \times 2 = 4\]
**Relational Database annotation: Fuzzy Databases**

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>Jane</td>
<td>Physics</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>History</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likes</th>
<th>Student</th>
<th>Course</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>Algebra</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>Literature</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>


(Jane, Algebra): $0.6 \times 0.5 = 0.3$
Semirings

(Green et al. 07):

- Domains of annotations are \textit{commutative semirings}.
- Typical example: \textit{natural numbers}
- $\mathcal{K} = \langle K, +, \times, 0, 1 \rangle$
Semirings

(Green et al. 07):

- Domains of annotations are **commutative semirings**.
- Typical example: **natural numbers**
- $\mathcal{K} = \langle K, +, \times, 0, 1 \rangle$

More examples:

- **Comments**: $\langle \{c_1, c_2, c_3, \ldots \}, \cup, \cup, \emptyset, \emptyset \rangle$
- **Belief**: $\langle x, y, z, \ldots, \cup, \cap, \emptyset, \emptyset \rangle$
- **Fuzzy Databases**: $\langle [0, 1], \max, \times, 0, 1 \rangle$
Semirings

(Green et al. 07):

- Domains of annotations are **commutative semirings**.
- Typical example: **natural numbers**
- \( K = \langle K, +, \times, 0, 1 \rangle \)

For query evaluation (positive relational algebra):

- Joins we **Multiply** the annotations
- Unions we **Add** the annotations
We study query containment in annotated databases
What is so **special** about containment?

- Not the same as Set Semantics
- **Varies** depending on the annotation domain
- **Open Problems** (Bag Semantics)
What is so special about containment?

- Not the same as Set Semantics
- Varies depending on the annotation domain
- Open Problems (Bag Semantics)

\[
Q_1 := \exists u \exists v, \exists w \text{ Takes}(u, v), \text{ Takes}(u, w)
\]
\[
Q_2 := \exists u \exists v \text{ Takes}(u, v)
\]
What is so **special** about containment?

- Not the same as Set Semantics
- **Varies** depending on the annotation domain
- **Open Problems** (Bag Semantics)

\[
Q_1 := \exists u \exists v, \exists w \text{ Takes}(u, v), \text{ Takes}(u, w)
\]

\[
Q_2 := \exists u \exists v \text{ Takes}(u, v)
\]

**Q₁ is contained in Q₂ under Set Semantics**

**Q₁ is not contained in Q₂ under Bag Semantics**
What is so special about containment?

- Not the same as Set Semantics
- Varies depending on the annotation domain
- Open Problems (Bag Semantics)

\[ Q_1 := \exists u \exists v, \exists w \text{ } Takes(u, v), Takes(u, w) \]
\[ Q_2 := \exists u \exists v \text{ } Takes(u, v) \]

\textbf{Q2 is contained in Q1 under Set Semantics or Bag Semantics}

\textbf{Q2 is not contained in Q1 over fuzzy databases}
Previous Work has focused on particular semirings

- Bag Semantics
- Probabilistic Databases
- Various semirings for provenance

But new applications may use new semirings

We focus on classes of semirings
Contributions

- Identify several classes of semirings for annotation with decision procedures for checking: containment of CQs and UCQs.

- Generalize previous work

- Some results by known techniques (homomorphisms)
- Others using new machinery, based on
  - Relationships between queries and polynomials
  - Small model properties
Outline

- Formalization of $\mathcal{K}$-containment
- Some results in the paper
Outline

- Formalization of $\mathcal{K}$-containment
- Some results in the paper
Query Evaluation on annotated databases

Bag Semantics: $\langle \mathbb{N}, +, \times \rangle$

$$Q := \exists u, \exists v, \exists w \ Takes(u, v), Takes(u, w)$$

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>P</td>
<td>1</td>
</tr>
</tbody>
</table>
Query Evaluation on annotated databases

Bag Semantics: $\langle \mathbb{N}, +, \times \rangle$

- For each homomorphism $h$ from $Q$ to $I$:
  1. Compute the annotation of $h(Q)$
  2. Sum over all homomorphisms.

$$Q := \exists u, \exists v, \exists w \text{Takes}(u, v), \text{Takes}(u, w)$$

<table>
<thead>
<tr>
<th>$I$: Takes</th>
<th>Student</th>
<th>Course</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>A</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>J</td>
<td>P</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
Query Evaluation on annotated databases

Bag Semantics: $\langle \mathbb{N}, +, \times \rangle$

For each homomorphism $h$ from $Q$ to $I$:
1. Compute the annotation of $h(Q)$
2. Sum over all homomorphisms.

$Q := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{ Takes}(u, w)$
$h(Q) := \text{ Takes}(J, A), \text{ Takes}(J, P)$

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>P</td>
<td>1</td>
</tr>
</tbody>
</table>

$Q_1(I) = 2 \cdot 1$
Query Evaluation on annotated databases

Bag Semantics: \( \langle \mathbb{N}, +, \times \rangle \)

- For each homomorphism \( h \) from \( Q \) to \( I \):
  1. Compute the annotation of \( h(Q) \)
  2. Sum over all homomorphisms.

\[
Q := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{ Takes}(u, w)
\]

\[
h(Q) := \text{ Takes}(J, P), \text{ Takes}(J, A)
\]

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>P</td>
<td>1</td>
</tr>
</tbody>
</table>

\[Q_1(I) = 2 \cdot 1 + 1 \cdot 2\]
Query Evaluation on annotated databases

Bag Semantics: $\langle \mathbb{N}, +, \times \rangle$

- For each homomorphism $h$ from $Q$ to $I$:
  1. Compute the annotation of $h(Q)$
  2. Sum over all homomorphisms.

\[
Q := \exists u, \exists v, \exists w \ \text{Takes}(u, v), \ \text{Takes}(u, w)
\]

\[
h(Q) := \text{Takes}(J, A), \ \text{Takes}(J, A)
\]

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td>J</td>
<td>P</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
Q_1(I) = 2 \cdot 1 + 1 \cdot 2 + 2 \cdot 2
\]
Query Evaluation on annotated databases

Bag Semantics: $\langle \mathbb{N}, +, \times \rangle$

For each homomorphism $h$ from $Q$ to $I$:
1. Compute the annotation of $h(Q)$
2. Sum over all homomorphisms.

$$Q := \exists u, \exists v, \exists w \; Takes(u, v), Takes(u, w)$$

$$h(Q) := Takes(J, P), Takes(J, P)$$

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>P</td>
<td>1</td>
</tr>
</tbody>
</table>

$$Q_1(I) = 2 \cdot 1 + 1 \cdot 2 + 2 \cdot 2 + 1 \cdot 1 = 9$$
Fuzzy Databases: $\langle [0, 1], \text{max}, \times \rangle$

- For each homomorphism $h$ from $Q$ to $I$:
  1. Compute the annotation of $h(Q)$
  2. Sum over all homomorphisms.

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$:</td>
<td>$J$</td>
<td>$A$</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>$J$</td>
<td>$P$</td>
<td>0.3</td>
</tr>
</tbody>
</table>

$Q_1(I) = \max (0.7 \times 0.3, 0.3 \times 0.7, 0.7 \times 0.7, 0.3 \times 0.3) = 0.49$
Query Containment over Annotated Databases

- Semirings with partial order $\preceq_K$
- For Bag Semantics, Fuzzy databases we use the order $\leq$
- For comments, belief, provenance we use order $\subseteq$
  \[
  \{\text{Wants TA}\} \subseteq \{\text{Top Mark, Wants TA}\}
  \]
Query Containment over Annotated Databases

- Semirings with partial order $\leq_K$

- For Bag Semantics, Fuzzy databases we use the order $\leq$

- For comments, belief, provenance we use order $\subseteq$:
  \[ \{\text{Wants TA}\} \subseteq \{\text{Top Mark, Wants TA}\} \]

Definition of containment (boolean queries):

\[ Q_1 \text{ is } K\text{-contained in } Q_2 \iff Q_1(I) \leq_K Q_2(I), \text{ for all instances } I \]

- Write $Q_1 \subseteq_K Q_2$
Outline

- Formalization of $\mathcal{K}$-containment
- Some results in the paper
Previous Work

Set semantics: $\langle \{0, 1\}, \lor, \land \rangle$
$Q_1 \subseteq Q_2$ iff homomorphism from $Q_2$ to $Q_1$.

Set semantics: $\langle \{0, 1\}, \lor, \land \rangle$
Previous Work

Positive Boolean Algebra

\[ Q_1 \subseteq_{K} Q_2 \text{ iff homomorphism from } Q_2 \text{ to } Q_1. \]
Previous Work - Grahne et al. ’97

Distributive Lattices

$B$

PosBool
Previous Work - Grahne et al. ’97

Distributive Lattices

\( \mathcal{B} \)

PosBool

homomorphism
Previous Work

Distributive Lattices

\( \mathcal{N} \)

\( \mathbf{B} \)

\( \text{PosBool} \)

\( \text{homomorphism} \)
If surjective homomorphism from $Q_2$ to $Q_1$, then $Q_1 \subseteq \kappa Q_2$. 

Distributive Lattices

$\mathcal{N}$

$\mathcal{B}$

PosBool

homomorphism
If $Q_1 \subseteq_K Q_2$ then homomorphic covering from $Q_2$ to $Q_1$. 

Distributive Lattices

PosBool

homomorphism
Previous Work - Green '09

Lineage

\[ \text{Distributive Lattices} \]

\[ \mathcal{B} \]

\[ \text{PosBool} \]

\[ \text{homomorphism} \]
Previous Work - Green ’09

Why[X]

Why - Provenance

Lineage

Distributive Lattices

PosBool

B

homomorphism
Previous Work - Green ’09

Polynomials over variables $X$ (Provenance Polynomials)

Why[$X$]

Lineage

$\mathcal{N}[X]$
$Q_1 \subseteq_K Q_2 \iff$ homomorphic covering from $Q_2$ to $Q_1$. 
Previous Work - Green ’09

\[ \mathcal{N}[X] \]

\[ \text{Why}[X] \]

\text{surjective homomorphism}

\text{Lineage}

\text{homomorphically covering}

\text{Distributive Lattices}

\text{PosBool}

\text{homomorphism}
Previous Work - Green ’09

\[ \mathcal{N}[X] \]
bijective homomorphism

\[ \text{Why}[X] \]
surjective homomorphism

Lineage
homomorphic covering

Distributive Lattices

\[ \mathcal{B} \]
homomorphism

\[ \text{PosBool} \]
Summing up, we have:

- Different types of mappings (homomorphisms)
- For a semiring $\mathcal{K}$ they can be:
  - *Sufficient condition* for containment
  - *Necessary condition* for containment
  - *Decision procedure* for containment
Summing up, we have:

- Different types of mappings (homomorphisms)
- For a semiring $\mathcal{K}$ they can be:
  - *Sufficient condition* for containment
    
    \[
    \text{If mapping from } Q_2 \text{ to } Q_1 \text{ then } Q_1 \subseteq_{\mathcal{K}} Q_2
    \]
  - *Necessary condition* for containment
  - *Decision procedure* for containment
Summing up, we have:

- Different types of mappings (homomorphisms)
- For a semiring $\mathcal{K}$ they can be:
  - *Sufficient condition* for containment
  - *Necessary condition* for containment

If $Q_1 \subseteq_{\mathcal{K}} Q_2$ then mapping from $Q_2$ to $Q_1$

- *Decision procedure* for containment
Summing up, we have:

- Different types of mappings (homomorphisms)
- For a semiring $\mathcal{K}$ they can be:
  - Sufficient condition for containment
  - Necessary condition for containment
  - Decision procedure for containment

\[ Q_1 \subseteq_{\mathcal{K}} Q_2 \text{ iff mapping from } Q_2 \text{ to } Q_1 \]
We fully characterize the universe of semirings
We fully characterize the universe of semirings

- Axiomatize classes of semirings for which different type of mappings are sufficient, or necessary conditions for $\mathcal{K}$-containment of CQ's

- Several classes for which $\mathcal{K}$-containment is decidable
We fully characterize the universe of semirings

- **Axiomatize classes of semirings** for which different type of mappings are sufficient, or necessary conditions for $\mathcal{K}$-containment of CQ’s

- Several classes for which $\mathcal{K}$-containment is **decidable**

- Generalize to **Unions of CQs**
We fully characterize the universe of semirings

- Axiomatize classes of semirings for which different type of mappings are sufficient, or necessary conditions for $K$-containment of CQ’s

- Several classes for which $K$-containment is decidable

- Generalize to Unions of CQs

- Additional decision procedures for $K$-containment
Outline

- Formalization of $\mathcal{K}$-containment
- Some results in the paper
- Results for homomorphisms
- Results for homomorphic covering...
  and a relevant class of polynomials
Containment of CQ's for set semantics

- Model set semantics as $B = \langle \{0, 1\}, \lor, \land, 0, 1 \rangle$

$Q_1$ is $B$-contained in $Q_2$ iff there is a homomorphism from $Q_2$ to $Q_1$
Containment of CQ's for set semantics

- Model set semantics as $B = \langle \{0, 1\}, \lor, \land, 0, 1 \rangle$

$Q_1$ is $B$-contained in $Q_2$ iff there is a homomorphism from $Q_2$ to $Q_1$

Is this true for any other semiring?
Many semirings behave as set semantics

- Boolean Algebra
- Event tables
- Type A systems (Ioannidis et al. 95)
- Distributive lattices
Many semirings behave as set semantics

- Boolean Algebra
- Event tables
- Type A systems (Ioannidis et al. 95)
- Distributive lattices

Can we characterize all semirings with this behavior?
Yes we can

A semiring $\mathcal{K}$ is in $\mathcal{H}$ if

1. $a \times a = a$
2. $1 + a = 1$

for all $a \in \mathcal{K}$. 
A semiring $\mathcal{K}$ is in $\mathcal{H}$ if

1. $a \times a = a$
2. $1 + a = 1$

for all $a \in \mathcal{K}$.

Theorem

$\mathcal{H}$ captures precisely all semirings that behave as Set Semantics (wrt. containment of CQs)
A semiring $\mathcal{K}$ is in $\mathcal{H}$ if

1. $a \times a = a$
2. $1 + a = 1$

for all $a \in \mathcal{K}$.

Theorem

$\mathcal{H}$ captures precisely all semirings that behave as Set Semantics (wrt. containment of CQs)

If $\mathcal{K}$ is in $\mathcal{H}$ then

- Homomorphism is a decision procedure for $\mathcal{K}$-containment
A semiring $\mathcal{K}$ is in $\mathcal{H}$ if

1. $a \times a = a$
2. $1 + a = 1$

for all $a \in \mathcal{K}$.

**Theorem**

$\mathcal{H}$ captures precisely all semirings that behave as Set Semantics (wrt. containment of CQs)

If Homomorphism is a decision procedure for $\mathcal{K}$-containment

➤ Then $\mathcal{K}$ is in $\mathcal{H}$
Class $\mathcal{H}$

$\mathcal{N}[X]$

Why$[X]$

$\mathcal{N}$

Lineage

Homomorphism from $Q_2$ to $Q_1$, iff $Q_1 \subseteq_{K} Q_2$

Distributive Lattices

PosBool

$\mathcal{B}$

$a \times a = a$

$1 + a = 1$
Outline

- Formalization of $\mathcal{K}$-containment
- Some results in the paper
- Results for homomorphisms
- Results for homomorphic covering...
  and a relevant class of polynomials
Moving away from $\mathcal{H}$

Two options:

- Keep $a \times a = a$
- Keep $1 + a = 1$
Moving away from $\mathcal{H}$

Two options:

- Keep $a \times a = a$
- Keep $1 + a = 1$

Example:

- Lineage $\text{Lineage} = \langle \{x, y, z, w, \ldots \}, \cup, \uplus \rangle$
Semirings satisfying $a \times a = a$
Semirings satisfying $a \times a = a$

- Homomorphisms are not sufficient condition

\[ Q_1 := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{Likes}(u, w) \]
\[ Q_2 := \exists u, \exists v \text{ Takes}(u, v) \]

- Homomorphism from $Q_2$ to $Q_1$
- $Q_1$ is not Lineage-contained in $Q_2$
Semirings satisfying $a \times a = a$

- Homomorphisms are not sufficient condition

$$Q_1 := \exists u, \exists v, \exists w \ Takes(u, v), Likes(u, w)$$
$$Q_2 := \exists u, \exists v \ Takes(u, v)$$

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>Lineage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$:</td>
<td>J</td>
<td>A</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>P</td>
<td>x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likes</th>
<th>Student</th>
<th>Course</th>
<th>Lineage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J</td>
<td>A</td>
<td>y</td>
</tr>
</tbody>
</table>
Semirings satisfying $a \times a = a$

- Homomorphisms are not sufficient condition

\[ Q_1 := \exists u, \exists v, \exists w \quad \text{Takes}(u, v), \text{Likes}(u, w) \]
\[ Q_2 := \exists u, \exists v \quad \text{Takes}(u, v) \]

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>Lineage</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I: )</td>
<td>J</td>
<td>A</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>P</td>
<td>x</td>
</tr>
<tr>
<td>(L: )</td>
<td>J</td>
<td>A</td>
<td>y</td>
</tr>
</tbody>
</table>

- \(Q_1(I) = \{x, y\}\)
- \(Q_2(I) = \{x\}\)
We need a stricter notion of mapping

Idea:

- force both queries to target the same relations
Homomorphic Covering from $Q_1$ to $Q_2$

Intuition:
Cover each atom of $Q_2$ with a homomorphism from $Q_1$ to $Q_2$

$Q_1 := \exists u, \exists v, \exists w \ Takes(u, v), Likes(u, w)$

$Q_2 := \exists u, \exists v, \exists w \ Takes(u, v), Takes(u, w), Likes(u, w)$
Homomorphic Covering from $Q_1$ to $Q_2$

Intuition:
Cover each atom of $Q_2$ with a homomorphism from $Q_1$ to $Q_2$

\[
Q_1 := \exists u, \exists v, \exists w \; Takes(u, v), Likes(u, w)
\]

\[
Q_2 := \exists u, \exists v, \exists w \; Takes(u, v), Takes(u, w), Likes(u, w)
\]
Homomorphic Covering from $Q_1$ to $Q_2$

Intuition:
Cover each atom of $Q_2$ with a homomorphism from $Q_1$ to $Q_2$

\[
Q_1 := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{ Likes}(u, w)
\]
\[
Q_2 := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{ Takes}(u, w), \text{ Likes}(u, w)
\]
Homomorphic Covering from $Q_1$ to $Q_2$

Intuition:
Cover each atom of $Q_2$ with a homomorphism from $Q_1$ to $Q_2$

\[
Q_1 := \exists u, \exists v, \exists w \; Takes(u, v), Likes(u, w) \\
Q_2 := \exists u, \exists v, \exists w \; Takes(u, v), Takes(u, w), Likes(u, w)
\]
Homomorphic Covering from $Q_1$ to $Q_2$

Intuition:
Cover each atom of $Q_2$ with a homomorphism from $Q_1$ to $Q_2$

\[
Q_1 := \exists u, \exists v, \exists w \ Takes(u, v), Likes(u, w)
\]
\[
Q_2 := \exists u, \exists v, \exists w \ Takes(u, v), Takes(u, w), Likes(u, w)
\]
Homomorphic Covering from $Q_1$ to $Q_2$

Intuition:
Cover each atom of $Q_2$ with a homomorphism from $Q_1$ to $Q_2$

$Q_1 := \exists u, \exists v, \exists w \ Takes(u, v), Likes(u, w)$

$Q_2 := \exists u, \exists v, \exists w \ Takes(u, v), Takes(u, w), Likes(u, w)$

There is a homomorphic covering from $Q_1$ to $Q_2$
Homomorphic Covering from $Q_1$ to $Q_2$

Intuition:
Cover each atom of $Q_2$ with a homomorphism from $Q_1$ to $Q_2$

$Q_1 := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{ Likes}(u, w)$
$Q_2 := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{ Takes}(u, w), \text{ Likes}(u, w)$

There is a homomorphic covering from $Q_1$ to $Q_2$

$Q_3 := \exists u, \exists v \text{ Takes}(u, v)$
$Q_4 := \exists u, \exists v, \exists w \text{ Takes}(u, v), \text{ Likes}(u, w)$

There is no homomorphic covering from $Q_3$ to $Q_4$
We can now capture semirings satisfying $a \times a = a$

Let $\mathcal{K}$ be a semiring.

**Theorem**

*If $\mathcal{K}$ satisfies $a \times a = a$ then Homomorphic covering is a sufficient condition for $\mathcal{K}$-containment*
We can now capture semirings satisfying \( a \times a = a \)

Let \( \mathcal{K} \) be a semiring.

**Theorem**

If \( \mathcal{K} \) satisfies \( a \times a = a \)

- Then *Homomorphic covering* is a *sufficient condition* for \( \mathcal{K} \)-containment

If *Homomorphic covering* is a *sufficient condition* for \( \mathcal{K} \)-containment

- Then, \( \mathcal{K} \) satisfies \( a \times a = a \)
We can now capture semirings satisfying $a \times a = a$

Let $\mathcal{K}$ be a semiring.

**Theorem**

If $\mathcal{K}$ satisfies $a \times a = a$

- Then Homomorphic covering is a **sufficient condition** for $\mathcal{K}$-containment

If Homomorphic covering is a **sufficient condition** for $\mathcal{K}$-containment

- Then, $\mathcal{K}$ satisfies $a \times a = a$

$a \times a = a$ captures homomorphic covering, as **sufficient condition**.
Class $\mathcal{H}$

If homomorphically covering from $Q_2$ to $Q_1$, then $Q_1 \subseteq_{\mathcal{K}} Q_2$

$\mathcal{N}[X]$

Lineage

$\mathcal{N}$

$\mathcal{B}$

$\mathcal{B}$

$\mathcal{N}$

Why $[X]$

$a \times a = a$
semirings $\mathcal{K}$ for which homomorphic covering is a necessary condition for $\mathcal{K}$-containment?

- Bag Semantics $\mathbb{N}$ should belong to this class.
- We axiomatize this class
- By abstracting query evaluation into polynomials.
When annotating each tuple with a different variable:

Evaluation of queries correspond to polynomials

- We need to understand the structure of these polynomials
\[ Q_1 := \exists u, \exists v, \exists w \ T akes(u, v), \ T akes(u, w) \]

\[
\begin{array}{c|c|c|c}
\text{Takes} & \text{Student} & \text{Course} & \text{P} \\
\hline
J & A & & x \\
J & P & & y \\
\end{array}
\]

\[ Q_1(I) = x \cdot y + y \cdot x + x \cdot x + y \cdot y = x^2 + 2xy + y^2 \]
CQ-admissible polynomials

Obtained from evaluating a CQ over an instance annotated with (different) variables.
CQ-admissible polynomials

Obtained from evaluating a CQ over an instance annotated with (different) variables.

- Not every polynomial is CQ-admissible
- Only homogeneous polynomials are
Only Homogeneous Polynomials

\[ Q_1 \ := \ \exists u, \exists v, \exists w \ \text{Takes}(u, v), \ \text{Takes}(u, w) \]

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(J)</td>
<td>(A)</td>
<td></td>
<td>(x)</td>
</tr>
<tr>
<td>(J)</td>
<td>(P)</td>
<td></td>
<td>(y)</td>
</tr>
</tbody>
</table>

\[ Q_1(I) = x \cdot y + y \cdot x + x \cdot x + y \cdot y = x^2 + 2xy + y^2 \]
Only Homogeneous Polynomials

\[ Q_1 := \exists u, \exists v, \exists w \; \text{Takes}(u, v), \text{Takes}(u, w) \]

<table>
<thead>
<tr>
<th>Takes</th>
<th>Student</th>
<th>Course</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J )</td>
<td>( J )</td>
<td>( A )</td>
<td>( x )</td>
</tr>
<tr>
<td>( J )</td>
<td>( J )</td>
<td>( P )</td>
<td>( y )</td>
</tr>
</tbody>
</table>

\[ Q_1(I) = x \cdot y + y \cdot x + x \cdot x + y \cdot y = x^2 + 2xy + y^2 \]

- Only homogeneous polynomials
- Precise definition is more technical
CQ-admissible polynomials

Obtained from evaluating a CQ over an instance annotated with (different) variables.

- Not every polynomial is CQ-admissible
- Only homogeneous polynomials are
- Every polynomial is UCQ-admissible
CQ-admissible polynomials

Obtained from evaluating a CQ over an instance annotated with (different) variables.

- Not every polynomial is CQ-admissible
- Only homogeneous polynomials are
- Every polynomial is UCQ-admissible

In the paper:

Syntactic characterization of CQ-admissible polynomials.
Homomorphic covering as necessary condition

Using CQ-admissible polynomials we define a class $C$ of semirings, such that:

**Theorem**

$C$ captures homomorphic coverings as a necessary condition

Note that $C$ contains Bag Semantics, but not Set Semantics.
And obtain a class where containment is decidable

The following are equivalent:

**Theorem**

- $\mathcal{K}$ belongs to $\mathcal{C}$ and satisfies $a \times a = a$
- $Q_1 \subseteq_{\mathcal{K}} Q_2$ iff homomorphic covering from $Q_2$ to $Q_1$

Gives us a large class of semirings where $\mathcal{K}$-containment is decidable
And obtain a class where containment is decidable

\[ N \subseteq_X B \]

Class \( C \)

Why \([X]\)

Homomorphic covering from \( Q_2 \) to \( Q_1 \), iff \( Q_1 \subseteq_K Q_2 \)

Lineage

\[ a \times a = a \]

\( \mathcal{N} \)
Also in the paper

- Similar theorems for surjective homomorphism, injective homomorphisms and bijective homomorphisms
- Extension to UCQs
- Complete Descriptions of CQs and UCQs
- Small model property and new procedures for semirings satisfying

\[ a + a = a \]
Future work

- Well behaved Semirings:
  \[ a + a = a \]

- Containment of \textit{CQ-admissible} polynomials over various semirings

- Views over annotated databases