Ontology-Based Query Answering for Probabilistic Temporal Data

Stream Reasoning Workshop, Linköping, 2019-04-17
Ontology-Based Query Answering

- Data coming from various sources
- Ontology defines background knowledge
- Queries answered wrt. augmented view
<table>
<thead>
<tr>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data (ABox)</strong></td>
</tr>
<tr>
<td>Male(peter) hasChild(peter, tom) worksAt(peter, TUDresden)</td>
</tr>
<tr>
<td><strong>Ontology (TBox)</strong></td>
</tr>
<tr>
<td>Father $\equiv$ Male $\sqcap$ $\exists$child. $\top$</td>
</tr>
<tr>
<td><strong>Conjunctive Query (CQ)</strong></td>
</tr>
<tr>
<td>$q(x) \leftarrow \exists y.\text{worksAt}(y, x) \wedge \text{Father}(y)$</td>
</tr>
</tbody>
</table>
Motivating Scenario

Application:
Ontology-based hypertension management in smartphone app
Motivating Scenario

Application:
Ontology-based hypertension management in smartphone app

- Sensor measures blood-pressure of patient
- Motion sensors indicate user activity
  - Walking
  - Cycling
  - Sitting
- Context + medical information
- Medical ontology as background knowledge
# Motivating Scenario

## Properties of Scenario

- History of observations relevant
  - Development of blood pressure
  - Recent activity

- A lot of observations *probabilistic* in nature
  - Uncertain measurements of sensor
  - Information inferred from motion sensor

⇒ Requires to handle data that are *temporal*
⇒ Requires to handle data that are *probabilistic*
**Temporal Probabilistic OBQA**

Sequence of Probabilistic Data

Augmented View on Data
Temporal Knowledge Bases

- Ontology + Sequence of datasets (ABoxes)
- Signature divided in two parts:
  - **Rigid** names
    - Interpretation independent of time
    - e.g. gender, has-parent relation,
  - **Non-rigid** names
    - Interpretation may change over time
    - e.g. blood pressure level, user activity
Example

### Sequence of ABoxes

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><code>bP(p, b)</code></td>
<td></td>
<td><code>High(b)</code></td>
<td></td>
<td><code>High(b)</code></td>
<td></td>
</tr>
</tbody>
</table>

### TBox

```
HighBloodPressurePatient ≡ ∃bloodPressure.High
```

### Rigid Names

- `bloodPressure (bP)`

⇒ `HighBloodPressurePatient(p) at 2 and 4`
Temporal Queries

- Well-investigated query language for temporal KBs
- Combine CQs with *Linear Temporal Logic* (LTL)

**Temporal queries (TQs)**

<table>
<thead>
<tr>
<th>(\exists \vec{x}. Q(\vec{x}, \vec{y}))</th>
<th>(\neg q)</th>
<th>(q_1 \land q_2)</th>
<th>(q_1 \lor q_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Box q)</td>
<td>(\Box \neg q)</td>
<td>(\Diamond q)</td>
<td>(\Diamond \neg q)</td>
</tr>
<tr>
<td>(\square q)</td>
<td>(\square \neg q)</td>
<td>(q_1 \mathcal{U} q_2)</td>
<td>(q_1 \mathcal{S} q_2)</td>
</tr>
</tbody>
</table>
Example Temporal Queries

Twice HighBloodPressurePatient within last 5 time units:

\[ q(x) \leftarrow \bigcirc^{-5} (\text{HBPP}(x) \land \bigtriangleup \text{HBPP}(x)) \]
Probabilistic OBQA

- Probabilistic ABoxes: Based on Probabilistic Databases
- Simplest approach: Assign probabilities to ABox axioms

| Exercise (patient): 0.6 | High Blood Pressure (patient): 0.8 |

- Define probability measure over possible worlds
- Assume statistical independence
  - More advanced models assign formulae over statistical variables
  ⇒ Complexities presented here still apply
## Temporal Probabilistic KBs

**Temporal Probabilistic ABox:**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$BP(p, b)$</td>
<td>1</td>
<td>High(b): 0.7</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
# Temporal Probabilistic KBs

## Temporal Probabilistic ABox:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BP(p, b)</td>
<td>High(b): 0.7</td>
<td></td>
<td>High(b): 0.9</td>
<td></td>
<td>High(b): 0.6</td>
</tr>
</tbody>
</table>

## Possible worlds:

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BP(p, b)</td>
<td>High(b)</td>
<td>High(b)</td>
<td>High(b)</td>
<td>High(b)</td>
<td>0.378</td>
</tr>
<tr>
<td>BP(p, b)</td>
<td>High(b)</td>
<td>High(b)</td>
<td></td>
<td></td>
<td>0.162</td>
</tr>
<tr>
<td>BP(p, b)</td>
<td>High(b)</td>
<td></td>
<td>High(b)</td>
<td></td>
<td>0.042</td>
</tr>
<tr>
<td>BP(p, b)</td>
<td>High(b)</td>
<td></td>
<td></td>
<td></td>
<td>0.018</td>
</tr>
<tr>
<td>BP(p, b)</td>
<td></td>
<td>High(b)</td>
<td>High(b)</td>
<td></td>
<td>0.252</td>
</tr>
<tr>
<td>BP(p, b)</td>
<td></td>
<td>High(b)</td>
<td></td>
<td></td>
<td>0.108</td>
</tr>
<tr>
<td>BP(p, b)</td>
<td></td>
<td></td>
<td>High(b)</td>
<td></td>
<td>0.028</td>
</tr>
<tr>
<td>BP(p, b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.012</td>
</tr>
</tbody>
</table>
Temporal Probabilistic Queries

Temporal Probabilistic Query Language

Conjunctive Queries + LTL operators + prob. operators

\[ P_{>p} Q \quad P_{=}p Q \quad P_{<p} Q \]

Example

\[ q(x) \leftarrow \bigcirc^{-10} \left( P_{<0.2}(\text{Exercising}(x)) \cup P_{>0.8}(\text{HighBPP}(x)) \right) \]

- During the **last 10 time units**, 
- patient was with **low probability** exercising 
- until with **high probability** he had high blood pressure
Complexities

- Classical
  - \text{ALCOIQ} → \text{ALCOIQ}
  - \text{ALCOIQ} → \text{ALCOIQ}

- No Rigid Roles
  - \text{ALCOIQ} → \text{ALCOIQ}
  - \text{ALCOIQ} → \text{ALCOIQ}

- Rigid Roles
  - \text{ALCOIQ} → \text{ALCOIQ}
  - \text{ALCOIQ} → \text{ALCOIQ}

- decidable
  - \text{ALCOIQ} → \text{SHIQ}
  - \text{SHIQ} → \text{ALCO}
  - \text{ALCO} → \text{ALCI}

- 2ExpTime
  - \text{ALC} → \text{SHQ}

- ExpTime
  - \text{EL}
  - \text{EL(data)}

- ExSpace
  - \text{EL(pos)}
  - \text{EL(pos,data)}

- P
  - \text{PL(pos)}

- NP
  - \text{PL(data)}

- PP
  - \text{PL(pos)}

- PP^NP - P^P
  - \text{PL(pos, data)}

- \text{EL(pos, data)}

- \text{EL(pos)}

- \text{EL(data)}

- \text{EL(pos)}

- \text{EL(pos, data)}

- \text{EL(pos, data)}
Complexities

- Classical
  - $\text{ALCHOTIQ} : \text{ALCOIQ}$
  - $\text{ALCOI} \ldots SHOI$
  - $\text{SHOQ} : \text{ALCO}$
  - $\text{SHIQ} : \text{ALCI}$
  - $\text{ALC} \ldots SHQ$
  - $\mathcal{E}L$
  - $\mathcal{E}L$($\text{data}$)

- No Rigid Roles
  - $\text{ALCHOTIQ} : \text{ALCOIQ}$
  - $\text{ALCOI} \ldots SHOI$
  - $\text{SHOQ} : \text{ALCO}$
  - $\text{SHIQ} : \text{ALCI}$
  - $\text{ALC} \ldots SHQ$
  - $\mathcal{E}L$($\text{pos}$)
  - $\mathcal{E}L$($\text{pos, data}$)

- Rigid Roles
  - $\text{ALCHOTIQ} : \text{ALCOIQ}$
  - $\text{ALCOI} \ldots SHOI$
  - $\text{SHOQ} : \text{ALCO}$
  - $\text{SHIQ} : \text{ALCI}$
  - $\text{ALC} \ldots SHQ$
  - $\emptyset \ldots \mathcal{E}LH$
  - $\mathcal{E}L$($\text{pos}$)
  - $\mathcal{E}L$($\text{pos, data}$)

Complexities:

- Decidable
- $2\text{ExpTime}$
- $\text{ExpTime}$
- $\text{ExpSpace}$
- $\text{NP}$
- $\text{P}$
- $\text{P}^{\text{NP}} - \text{P}^{\text{NP}}$
- $\mathcal{P}^{\text{P}}$
Complexities

classical

No Rigid Roles

Rigid Roles

\( \text{ALCHOIQ} \)

\( \text{ALCOIQ} \)

decidable

\( \text{ALCOIQ} \)

\( \text{ALCOIQ} \)

\( \text{ALCOIQ} \)

2\( \text{ExpTime} \)

\( \text{ExpTime} \)

\( \text{ExpSpace} \)

NP

P

\( \text{EL} \)

\( \text{EL} \)

\( \text{EL(pos)} \)

\( \text{EL(pos)} \)

\( \text{EL(data)} \)

\( \text{EL(pos, data)} \)

PP

PP

\( \text{P} \)

\( \text{P} \)

\( \text{P} \)

\( \text{P} \)
Complexities Probabilistic Temporal OBQA

- **Classical**
  - Decidable
  - $2\text{ExpTime}$
  - $\text{ExpTime}$
  - NP
  - P

- **No Rigid Roles**
  - Decidable
  - $\text{ExpSpace}$
  - $\mathbb{P}^\mathbb{NP}$, $\mathbb{P}^\mathbb{PP}$
  - $\mathbb{P}$

- **Rigid Roles**
  - Decidable
  - $\text{ExpSpace}$
  - $\mathbb{P}^\mathbb{NP}$, $\mathbb{P}^\mathbb{PP}$
  - $\mathbb{P}$

---

Ontology-Based Query Answering for Probabilistic Temporal Data
© Patrick Koopmann
Stream Reasoning Workshop, Linköping, 2019-04-17
### Main results

- Framework combining temporal and probabilistic OBQA
- \( \text{EXPSPACE} \)-hard, already without ontology
- Source of complexity: negation
  - Without, not much harder as probabilistic data access (PP/PP\(^{NP} \))

### Current Research

- Prototypical implementation
  - positive queries, \textit{DL-Lite}
- Extend also ontology language
  - prob.+temp. concept+axiom operators
  - prob. + temp. on concepts makes 2-\text{EXPSPACE}-hard