Extending NoHR for OWL 2 QL

Nuno Costa  Matthias Knorr  João Leite

NOVALINCS
LABORATORY FOR COMPUTER SCIENCE AND INFORMATICS

Universidade Nova de Lisboa
Motivation: OWA vs. CWA

- **Open World Assumption (OWA)**
  - Model taxonomic knowledge
  - Ontologies (in Description Logics (DL), such as $\mathcal{EL}$, $DL$-$\text{Lite}_R$)
  - Example: results of clinical tests

- **Closed World Assumption (CWA)**
  - Model defaults and exceptions
  - Non-monotonic rules well-suited
  - Example: patient’s medication

Integration for benefits of both approaches
Motivation: OWA vs. CWA

- **Open World Assumption (OWA)**
  - Model taxonomic knowledge
  - Ontologies (in Description Logics (DL), such as $\mathcal{EL}$, $DL$-Lite$_R$)
  - Example: results of clinical tests

- **Closed World Assumption (CWA)**
  - Model defaults and exceptions
  - Non-monotonic rules well-suited
  - Example: patient’s medication

Integration for benefits of both approaches
Motivation: OWA vs. CWA

- **Open World Assumption (OWA)**
  - Model taxonomic knowledge
  - Ontologies (in Description Logics (DL), such as $\mathcal{EL}$, $DL$-Lite$_R$)
  - Example: results of clinical tests

- **Closed World Assumption (CWA)**
  - Model defaults and exceptions
  - Non-monotonic rules well-suited
  - Example: patient’s medication

Integration for benefits of both approaches
Requirements for Integration

1. Flexible framework
   ▶ Expressive language, yet simple to use
   ▶ Full two-way interaction between ontologies and rules
   ▶ As little restrictions as possible

2. Low complexity
   ▶ Large amount of data (on the Web; in applications, e.g., patient records)
   ▶ Interactive response time on reasoning

3. Top-down querying
   ▶ Avoid up-front computation of the entire model
   ▶ Restrict computation to the relevant part
Requirements for Integration

1. Flexible framework
   - Expressive language, yet simple to use
   - Full two-way interaction between ontologies and rules
   - As little restrictions as possible

2. Low complexity
   - Large amount of data (on the Web; in applications, e.g., patient records)
   - Interactive response time on reasoning

3. Top-down querying
   - Avoid up-front computation of the entire model
   - Restrict computation to the relevant part
Requirements for Integration

1. Flexible framework
   ▶ Expressive language, yet simple to use
   ▶ Full two-way interaction between ontologies and rules
   ▶ As little restrictions as possible

2. Low complexity
   ▶ Large amount of data (on the Web; in applications, e.g., patient records)
   ▶ Interactive response time on reasoning

3. Top-down querying
   ▶ Avoid up-front computation of the entire model
   ▶ Restrict computation to the relevant part
1. Hybrid MKNF [Motik and Rosati, J. ACM 2010]
2. Its Well-Founded Semantics (WFS) [Knorr et al., AI 2011]
3. Top-down procedure \( \text{SLG}(\mathcal{O}) \) [Alferes et al., ACM TOCL 2013]
Motivation: Extension to QL

- Applications require DL language features (e.g., inverses) [Calvanese et al., 2011] not covered by OWL EL
- OWL QL based on $DL-Lite_R$ would serve
  - Covers basic DL languages, the entity relationship model, and basic UML class diagrams
  - Query-answering by rewriting queries by means of the ontology s.t. SQL engines can be used over the data
  - Very low data complexity
  - Tailored towards huge data sets
Motivation: Extension to QL

- Applications require DL language features (e.g., inverses) [Calvanese et al., 2011] not covered by OWL EL
- OWL QL based on $DL-Lite_R$ would serve
  - Covers basic DL languages, the entity relationship model, and basic UML class diagrams
  - Query-answering by rewriting queries by means of the ontology s.t. SQL engines can be used over the data
  - Very low data complexity
  - Tailored towards huge data sets
Problem

- Negation present in OWL QL requires classification of negated concepts
- Currently no classifier for OWL QL including negated concepts
- Naive adaptation inefficient due to large number of created axioms

Objective

Adapt NoHR to OWL QL

- Direct translation (no prior classification)
- Ensure identical derivation of ground queries
- Implement and evaluate its performance
Problem

- Negation present in OWL QL requires classification of negated concepts
- Currently no classifier for OWL QL including negated concepts
- Naive adaptation inefficient due to large number of created axioms

Objective

Adapt NoHR to OWL QL
- Direct translation (no prior classification)
- Ensure identical derivation of ground queries
- Implement and evaluate its performance
\[ DL-Lite_R \]

\[ B \rightarrow A \mid \exists Q \quad C \rightarrow B \mid \neg B \quad Q \rightarrow P \mid P^- \quad R \rightarrow Q \mid \neg Q \]

\( A \in N_C \) concept name, \( P \in N_R \) role name, and \( P^- \) its inverse

- GCIs \( B \sqsubseteq C \) and RIs \( Q \sqsubseteq R \)
- Standard DL semantics based on interpretations \( \mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I}) \)

\[ \exists \text{HasArtist}^- \sqsubseteq \text{Artist} \quad \text{Piece} \sqsubseteq \exists \text{HasArtist} \]
\[ \exists \text{HasComposed}^- \sqsubseteq \text{Piece} \quad \text{Artist} \sqsubseteq \neg \text{Piece} \]
\[ \text{HasComposed}^- \sqsubseteq \text{HasArtist} \]
\( DL-Lite_R \)

\[
B \rightarrow A \mid \exists Q \quad C \rightarrow B \mid \neg B \quad Q \rightarrow P \mid P^- \quad R \rightarrow Q \mid \neg Q
\]

\( A \in N_C \) concept name, \( P \in N_R \) role name, and \( P^- \) its inverse

- GCIs \( B \sqsubseteq C \) and RIs \( Q \sqsubseteq R \)
- Standard DL semantics based on interpretations \( \mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I}) \)

\[
\exists \text{HasArtist}^- \sqsubseteq \text{Artist} \quad \text{Piece} \sqsubseteq \exists \text{HasArtist} \\
\exists \text{HasComposed}^- \sqsubseteq \text{Piece} \quad \text{Artist} \sqsubseteq \neg \text{Piece} \\
\text{HasComposed}^- \sqsubseteq \text{HasArtist}
\]
**Direct Translation**

\[ \text{Piece} \sqsubseteq \exists \text{HasArtist} \] cannot be translated naively

- \[ \text{HasArtist}(x, y) \leftarrow \text{Piece}(x) \] would yield \( \text{HasArtist}(x, y) \) for any \( \text{Piece}(x) \) and \( y \)
- \[ \text{HasArtist}(x, c) \leftarrow \text{Piece}(x) \] would yield \( \text{HasArtist}(x, c) \) for any \( \text{Piece}(x) \) for the same \( c \)
- Skolemization would cause difficulties for termination

**Special predicates for domain and range**

\[ \text{DHasArtist}(x) \leftarrow \text{Piece}(x) \] with \( \text{DHasArtist} \) the domain of \( \text{HasArtist} \) (and \( R\text{HasArtist} \) its range)
Direct Translation

\[ \text{Piece} \sqsubseteq \exists \text{HasArtist} \text{ cannot be translated naively} \]

- \( \text{HasArtist}(x, y) \leftarrow \text{Piece}(x) \) would yield \( \text{HasArtist}(x, y) \) for any \( \text{Piece}(x) \) and \( y \)
- \( \text{HasArtist}(x, c) \leftarrow \text{Piece}(x) \) would yield \( \text{HasArtist}(x, c) \) for any \( \text{Piece}(x) \) for the same \( c \)
- Skolemization would cause difficulties for termination

Special predicates for domain and range

\( \text{DHasArtist}(x) \leftarrow \text{Piece}(x) \) with \( \text{DHasArtist} \) the domain of \( \text{HasArtist} \) (and \( \text{RHasArtist} \) its range)
Direct Translation (2)

- $DHasArtist(x) ← HasArtist(x, y)$ associating domains (and ranges) to binary atoms
- For inverses $\text{HasComposed}^\neg \subseteq \text{HasArtist}$, translate to
  
  $$HasArtist(x, y) ← \text{HasComposed}(y, x)$$

  also link both auxiliary predicates via

  $$DHasArtist(x) ← RHasComposed(x)$$

  $$RHasArtist(x) ← DHasComposed(x)$$
Direct Translation (2)

- \( D\text{HasArtist}(x) \leftarrow \text{HasArtist}(x, y) \) associating domains (and ranges) to binary atoms

- For inverses \( \text{HasComposed}^- \sqsubseteq \text{HasArtist} \), translate to

  \[
  \text{HasArtist}(x, y) \leftarrow \text{HasComposed}(y, x)
  \]

  also link both auxiliary predicates via

  \[
  D\text{HasArtist}(x) \leftarrow R\text{HasComposed}(x)
  \]

  \[
  R\text{HasArtist}(x) \leftarrow D\text{HasComposed}(x)
  \]
Graph Representation Including Negation

Nodes all general concepts and roles, edges GCIs and RIs (including, e.g., implicit contrapositives)

\[
\exists \text{HasComposed} \\
\text{Piece} \\
\neg \text{Artist} \\
\neg \exists \text{HasArtist} \\
\neg \exists \text{HasComposed} \\
\neg \exists \text{HasComposed} \\

\exists \text{HasComposed} \\
\neg \text{HasArtist} \\
\neg \text{HasComposed} \\
\neg \exists \text{HasArtist} \\
\neg \exists \text{HasComposed} \\
\neg \exists \text{HasComposed} \\

\neg \text{HasComposed} \\
\neg \exists \text{HasComposed} \\

\text{HasComposed irreflexive: } \exists \text{HasComposed} \sqsubseteq \neg \exists \text{HasComposed} \\

\]

Computing irreflexive roles and unsatisfiable roles and (atomic) concepts necessary
Nodes all general concepts and roles, edges GCIs and RIs (including, e.g., implicit contrapositves)

Graph Representation Including Negation

Has\text{Composed} \text{ irreflexive: } \exists \text{HasComposed} \sqsubseteq \neg \exists \text{HasComposed}^\text{\neg

Computing irreflexive roles and unsatisfiable roles and (atomic) concepts necessary
Results

- Sound and complete translation w.r.t. answering (ground) queries
- Data complexity in P
- Extension of classification on graphs to negated concepts a contribution in its own right
- Implementation as an alternative translator module in NoHR for OWL QL
Evaluation Settings

LUBM benchmark

▶ Small TBox
▶ Data generator for creating instance data of large sizes
▶ 14 test queries

Here:

▶ TBox slightly simplified to match the OWL profile(s)
▶ Three queries omitted whose results are affected by the simplifications
Evaluation: Preprocessing

Direct translation approach vs. classification-based – LUBM reduced to fit OWL QL and EL to compare NoHR QL and EL approaches

QL considerably faster (up to 80s for LUBM_{20}) – due to avoiding classification and a smaller rule file being created
Direct translation approach vs. classification-based – LUBM reduced to fit OWL QL and EL to compare NoHR QL and EL approaches

QL considerably faster (up to 80s for LUBM_{20}) – due to avoiding classification and a smaller rule file being created
Evaluation: Querying

11 queries of the LUBM queries tested, representatives shown

- Often interactive response time with slight advantage for EL ($q_5$)
- Few take a considerable amount of time
  - Some with slight advantage for OWL QL ($q_{14}$)
  - One with notable difference in favor of EL ($q_9$)
11 queries of the LUBM queries tested, representatives shown

▸ Often interactive response time with slight advantage for EL ($q_5$)
▸ Few take a considerable amount of time
  ▸ Some with slight advantage for OWL QL ($q_{14}$)
  ▸ One with notable difference in favor of EL ($q_9$)
749 subclass axioms, 1,486 class disjointness axioms and 20 inverse object properties in combination with non-monotonic rules

- Preprocessing very fast and only linearly increasing
- Four atomic queries in different levels of the hierarchy with interactive response time
- 4’ – query 4 without the other queries beforehand (tabling)
Conclusions

- NoHR extended to OWL 2 QL based on direct translation
- Theoretically sound and complete including novel extension of graph-based reasoning with negated concepts
- Evaluation results of implementation encouraging as all previously observed results (for EL) persist
- QL is even faster on pre-processing and only slightly slower on average when answering queries
Future Work

- Further comparisons to alternative versions for QL based on, e.g., ontop, Konclude
- OWL RL
- Paraconsistent Semantics