Combining Logic Programming with Description Logics and Machine Learning for the Semantic Web

Francesca A. Lisi
lisi@di.uniba.it

Floriana Esposito
esposito@di.uniba.it

Dipartimento di Informatica
Università degli Studi di Bari
Via Orabona, 4 - 70126 Bari - Italy

ALPSWS 08
Motivation

Acquiring and maintaining rules is a demanding task. Machine Learning can partially automate this task.

Learning Semantic Web rules
≈
Learning Datalog rules on top of OWL ontologies
≈
Learning Datalog rules by having OWL ontologies as BK

Combining LP with Description Logics and Machine Learning

Dr. Francesca A. Lisi
Overview

- Motivation
- Background
- Combining LP and DLs with DL+log
- Inducing SHIQ+log Rules with ILP
- Related work
- Conclusions and future work
LP and Description Logics

**DLs vs HCL**

- **Different expressive power** (Borgida, 1996)
  - No relations of arbitrary arity or arbitrary joins between relations in DLs
  - No exist. quant. in HCL
- **Different semantics** (Rosati, 2005)
  - OWA for DLs
  - CWA for HCL
- **Can they be combined?** Yes, but integration can be easily undecidable if unrestricted

Dr. Francesca A. Lisi
LP and Description Logics (2)

Hybrid DL-HCL KR systems

- **CARIN** (Levy & Rousset, 1998)
  - Any DL+HCL
  - Unsafe
  - Decidable for some simple DL (e.g., ALCNR)

- **AL-log** (Donini et al., 1998)
  - ALC+Datalog
  - Safe
  - Decidable

- **DL+log** (Rosati, 2006)
  - Any DL+ Datalog\(^\neg\)
  - Weakly-safe
  - Decidable for some v.e. DL (e.g., SHIQ)
LP and Machine Learning

Inductive Logic Programming

- Use of prior knowledge
- Use of Datalog as KR framework
- Use of Concept Learning notions
  - Generalization as search through a partially ordered space of hypotheses
LP and Machine Learning (2)

- Learning in Carin-ALN (Rouveirol & Ventos, 2000)
- Learning in AL-log (Lisi, 2008)
Overview

- Motivation
- Background
- Combining LP and DLs with DL+log
  - Syntax
  - Semantics
  - Reasoning
- Inducing SHIQ+log¬ Rules with ILP
- Related work
- Conclusions and future work
Combining LP & DLs with DL+log: syntax

DL+log KB = DL KB extended with Datalog\(^\lor\) rules

\[ p_1(X_1) \lor ... \lor p_n(X_n) \leftarrow r_1(Y_1), ..., r_m(Y_m), s_1(Z_1), ..., s_k(Z_k), \neg u_1(W_1), ..., \neg u_h(W_h) \]

satisfying the following properties

\* Datalog safeness: every variable occurring in a rule must appear in at least one of the atoms \( r_1(Y_1), ..., r_m(Y_m), s_1(Z_1), ..., s_k(Z_k) \)

\* DL weak safeness: every head variable of a rule must appear in at least one of the atoms \( r_1(Y_1), ..., r_m(Y_m) \)
Combining LP & DLs with DL+log: semantics

- **FOL-semantics**
  - OWA for both DL and Datalog predicates

- **NM-semantics**: extends stable model semantics of Datalog\(\neg \lor\)
  - OWA for DL-predicates
  - CWA for Datalog-predicates

- In both semantics, entailment can be reduced to satisfiability

- In Datalog\(\lor\), FOL-semantics equivalent to NM-semantics
Combining LP & DLs with DL+log: reasoning

- CQ answering can be reduced to satisfiability
- NM-satisfiability of DL+log KBs combines
  - Consistency in Datalog\(^\neg\neg\) : A Datalog\(^\neg\neg\) program is consistent if it has a stable model
  - Boolean CQ/UCQ containment problem in DLs: Given a DL-TBox T, a Boolean CQ \(Q_1\) and a Boolean UCQ \(Q_2\) over the alphabet of concept and role names, \(Q_1\) is contained in \(Q_2\) wrt T, denoted by \(T \models Q_1 \subseteq Q_2\), iff, for every model I of T, if \(Q_1\) is satisfied in I then \(Q_2\) is satisfied in I.

- The decidability of reasoning in DL+log depends on the decidability of the Boolean CQ/UCQ containment problem in DL
  - SHIQ+log = most powerful decidable instantiation of DL+log!
Overview

- Motivation
- Background
- Combining LP and DLs with DL+log

**Inducing SHIQ+log¬ Rules with ILP**
- The problem statement
- The hypothesis ordering
- The hypothesis coverage of observations

- Related work
- Conclusions and future work
Inducing SHIQ+log rules with ILP: the problem statement

- Learning rules from ontologies and relational data
  - Rules for defining new relations
  - Rules for defining new concepts/roles

- Scope of induction: discrimination/characterization

- ILP setting: learning from interpretations

- Language choice: SHIQ+log¬ (SHIQ+Datalog¬)
  - Hypothesis as linked and connected SHIQ+log¬ rules
  - NAF literal ¬p(X) transformed into not_p(X)
**Inducing SHIQ+log rules with ILP: the problem statement (2)**

[A1] \[ \text{RICH} \sqsubseteq \text{UNMARRIED} \sqsubseteq \exists \text{WANTS-TO-MARRY}.T \]

[R1] \[ \text{RICH}(X) \leftarrow \text{famous}(X), \neg \text{scientist}(X) \]

\[ L^{\text{happy}} \]
\[ \text{in } \{ \text{famous}/1, \text{RICH}/1, \text{WANTS-TO-MARRY}/2, \text{LIKES}/2 \} \]
\[ \text{happy}(X) \leftarrow \text{famous}(X), \text{WANTS-TO-MARRY}(Y,X) \]

\[ L^{\text{LONER}} \]
\[ \text{in } \{ \text{famous}/1, \text{scientist}/1, \text{UNMARRIED}/1 \} \]
\[ \text{LONER}(X) \leftarrow \neg \text{famous}(X) \]

UNMARRIED(Mary)
UNMARRIED(Joe)
famous(Mary)
famous(Paul)
famous(Joe)
LONER(Joe)
Inducing SHIQ+log rules with ILP: the hypothesis ordering

\begin{itemize}
  \item SHIQ+log\neg KB \mathcal{K}
  \item SHIQ+log\neg rules H_1, H_2 \in \mathcal{L}
  \item Skolem substitution \sigma for H_2 w.r.t. \{H_1\} \cup \mathcal{K}
\end{itemize}

H_1 subsumes H_2 w.r.t. \mathcal{K} iff there exists a ground substitution \theta for H_1 such that

\begin{itemize}
  \item head(H_1)\theta = head(H_2)\sigma
  \item \mathcal{K} \cup body(H_2)\sigma \models body(H_1)\theta
\end{itemize}

Generality order boils down to CQ answering!
Inducing SHIQ+log rules with ILP: the hypothesis ordering (2)

[A1] \( \text{RICH} \cap \text{UNMARRIED} \subseteq \exists \text{WANTS-TO-MARRY} \cdot \top \)

[R1] \( \text{RICH}(X) \leftarrow \text{famous}(X), \neg \text{scientist}(X) \)

\( \chi_1^{\text{happy}} = \text{happy}(A) \leftarrow \text{RICH}(A) \)

\( \chi_2^{\text{happy}} = \text{happy}(X) \leftarrow \text{famous}(X) \)

\( \chi_1^{\text{happy}} \not\succ_{\mathcal{K}} \chi_2^{\text{happy}} \)

\( \chi_2^{\text{happy}} \not\succ_{\mathcal{K}} \chi_1^{\text{happy}} \)
Inducing SHIQ+log rules with ILP: the coverage relations

- SHIQ+log¬ KB $\mathcal{K}$
- SHIQ+log¬ rule $H \in \mathcal{L}$
- Observation $o_i = (p(a_i), F_i)$ where:
  - $a_i$ is an individual
  - $F_i$ is a set of ground Datalog facts

$H$ covers $o_i$ under interpretations w.r.t. $\mathcal{K}$ iff $\mathcal{K} \cup F_i \cup H \models p(a_i)$

Coverage boils down to CQ answering!
Inducing SHIQ+log rules with ILP: the coverage relations (2)

[A1] $\text{RICH} \land \text{UNMARRIED} \sqsubseteq \exists Y \text{ WANTS-TO-MARRY}$. $T$

[R1] $\text{RICH}(X) \leftarrow \text{famous}(X), \neg \text{scientist}(X)$

$H = \text{happy}(X) \leftarrow \text{famous}(X), \text{WANTS-TO-MARRY}(Y,X)$

covers $o_{\text{Mary}} = (\text{happy}(\text{Mary}), F_{\text{Mary}})$ because

$K \cup F_{\text{Mary}} \cup H \models \text{happy}(\text{Mary})$. 
Overview

- Motivation
- Background
- Combining LP and DLs with DL+log
- Inducing SHIQ+log¬ Rules with ILP
- Related work
- Conclusions and future work
Related work

<table>
<thead>
<tr>
<th>prior knowledge</th>
<th>Learning in CarIN-(\mathcal{ALN}) [24]</th>
<th>Learning in (\mathcal{AL})-log [15]</th>
<th>Learning in (\mathcal{SHIQ}+\log^-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ontology lang.</td>
<td>(\mathcal{ALN}) KB (\mathcal{ALN})</td>
<td>(\mathcal{AL})-log KB (\mathcal{AL})</td>
<td>(\mathcal{SHIQ}+\log^-) KB (\mathcal{SHIQ})</td>
</tr>
<tr>
<td>rule lang.</td>
<td>Horn clauses</td>
<td>(\mathcal{AL})-log clauses (\mathcal{AL})-log</td>
<td>(\mathcal{SHIQ}+\log^-) clauses (\mathcal{SHIQ}+\log^-)</td>
</tr>
<tr>
<td>hypothesis lang.</td>
<td>CarIN-(\mathcal{ALN}) non-recursive rules</td>
<td>constrained (\mathcal{AL})-log clauses (\mathcal{AL})-log</td>
<td>(\mathcal{SHIQ}+\log^-) non-recursive rules (\mathcal{SHIQ}+\log^-)</td>
</tr>
<tr>
<td>target predicate</td>
<td>Horn literal</td>
<td>(\mathcal{AL})-log literal (\mathcal{AL})-log</td>
<td>(\mathcal{SHIQ}/\mathcal{AL})-log literal (\mathcal{SHIQ}/\mathcal{AL})-log</td>
</tr>
<tr>
<td>observations</td>
<td>interpretations/implications</td>
<td>interpretations/predictive/implications</td>
<td>interpretations/predictive/implications</td>
</tr>
<tr>
<td>induction</td>
<td>predictive</td>
<td>predictive/predictive</td>
<td>predictive/predictive</td>
</tr>
<tr>
<td>order</td>
<td>CarIN-(\mathcal{ALN}) query answering</td>
<td>(\mathcal{AL})-log query answering</td>
<td>(\mathcal{SHIQ}+\log^-) query answering</td>
</tr>
<tr>
<td>coverage test</td>
<td>no</td>
<td>downward</td>
<td>no</td>
</tr>
<tr>
<td>ref. operators</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>implementation</td>
<td>partially</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>application</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Overview

- Motivation
- Background
- Combining LP and DLs with DL+log
- Inducing SHIQ+log¬ Rules with ILP
- Related work
- Conclusions and future work
Conclusions

- ILP can help learning Semantic Web rules

- DL+log is good for representing Semantic Web rules
  - Parametric wrt the DL part
  - Decidable for many DLs, notably SHIQ

- ILP in SHIQ+log¬ is feasible
  - Decidable coverage and generality relations
  - Valid for any decidable instantiation of DL+log with Datalog¬
Future work

- To study the impact of having Datalog\(\neg\lor\) both in the language of hypotheses and in the language for the BK
  - Nonmonotonic features to deal with incomplete knowledge

- To define ILP algorithms starting from the ingredients identified in this paper.

- To apply these algorithms to use cases for Semantic Web rules
  - See SWAP’08 for an application to ontology evolution