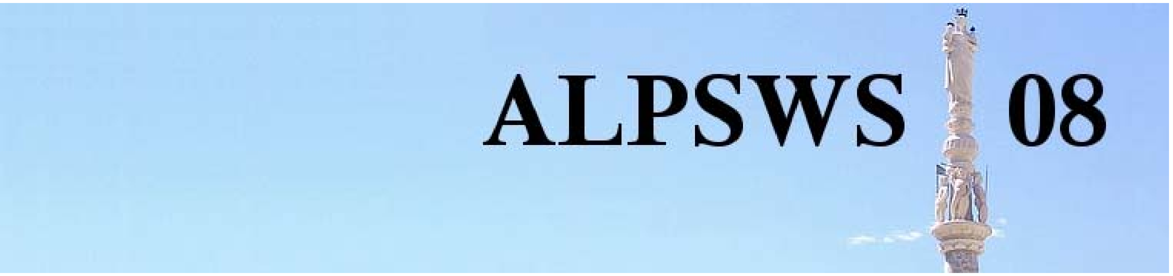
Combining Logic Programming with Description Logics and Machine Learning for the Semantic Web

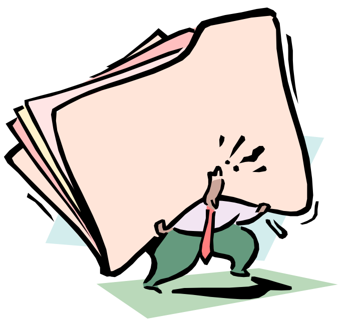
Francesca A. Lisi
lisi@di.uniba.it

Floriana Esposito
esposito@di.uniba.it

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



Motivation



Acquiring and mantaining rules is a demanding task

Machine Learning can partially automate this task

Learning Semantic Web rules

≈

Learning Datalog rules on top of OWL ontologies

≈

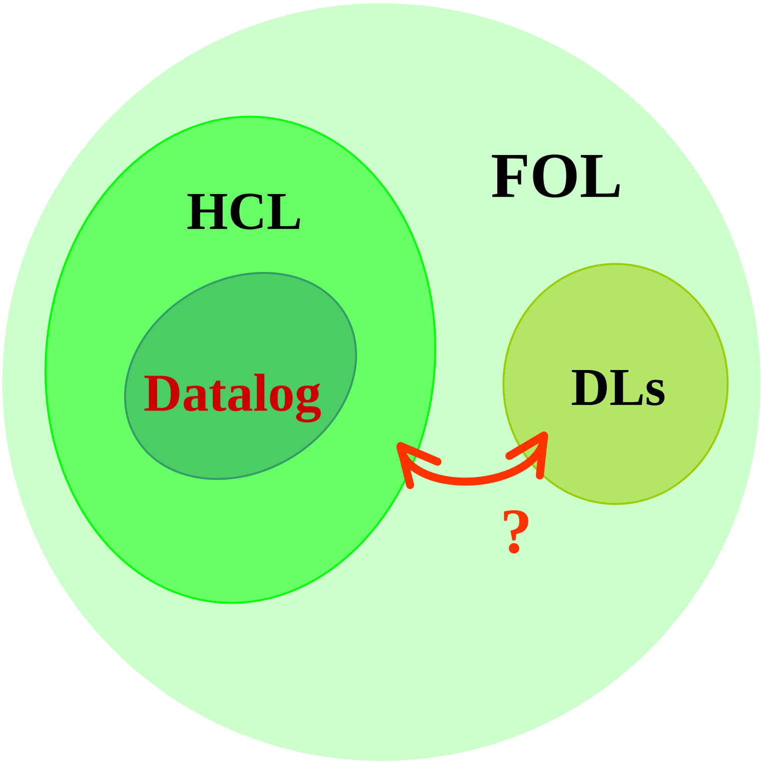
Learning Datalog rules by having OWL ontologies as BK

Combining LP with Description Logics and Machine Learning

Overview

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

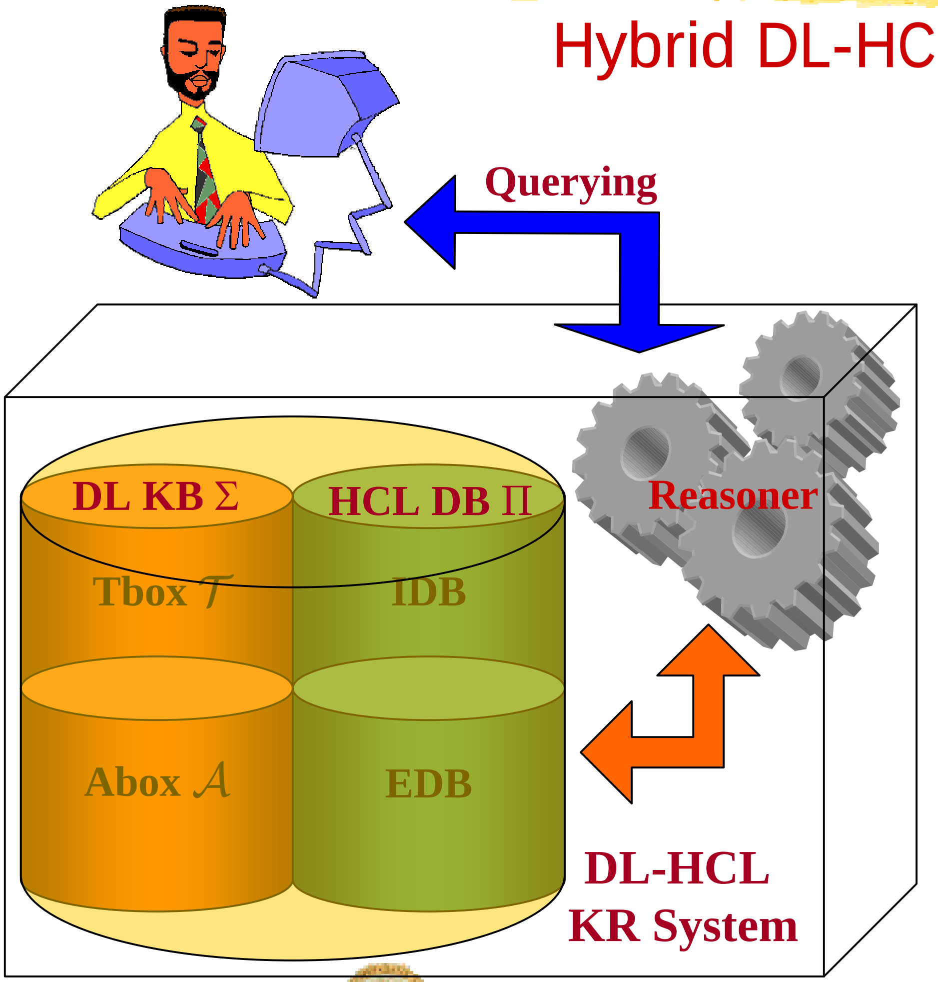
LP and Description Logics



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

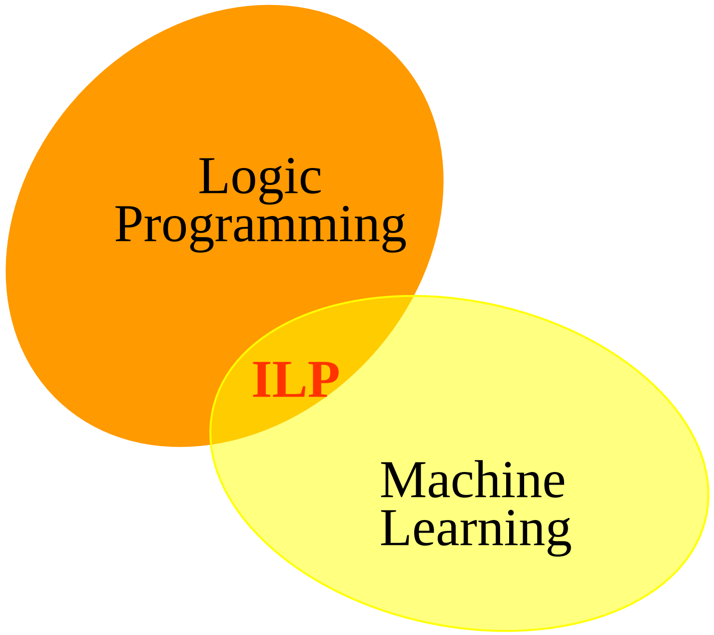
LP and Description Logics (2)



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$^{\neg\vee}$
  - Weakly-safe
  - Decidable for some v.e. DL (e.g., SHIQ)

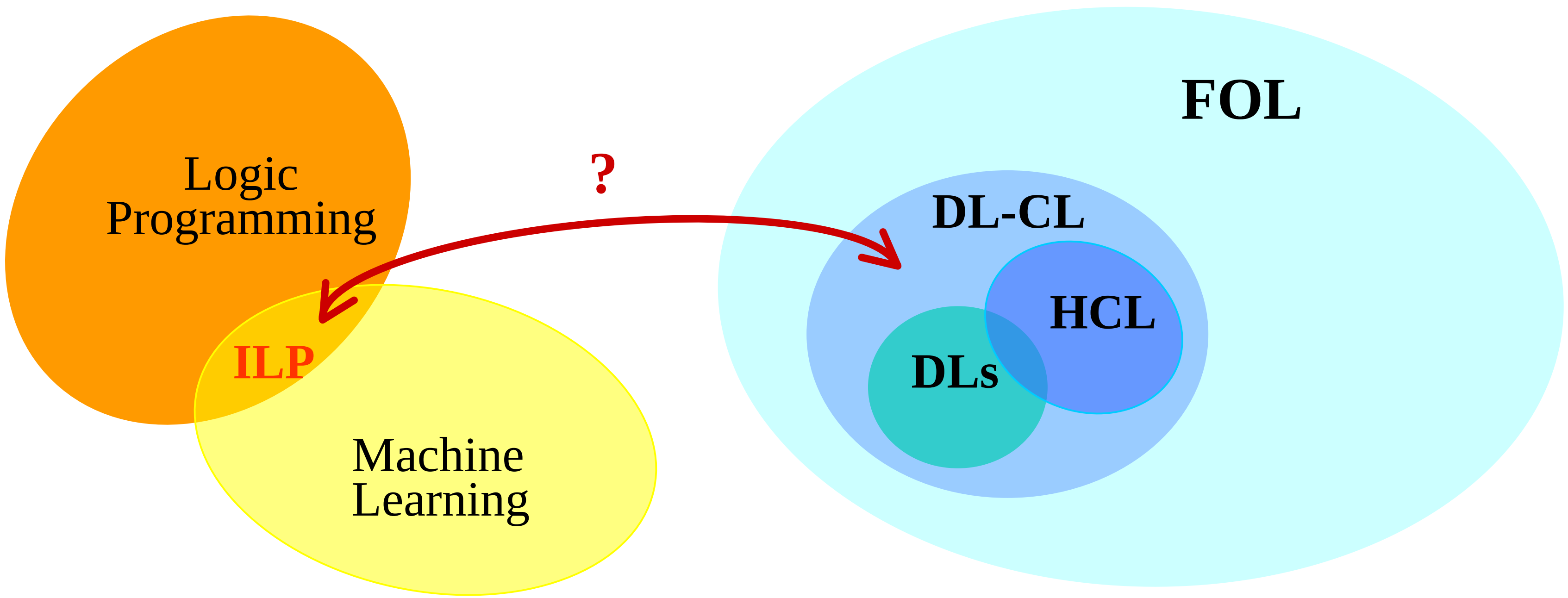
LP and Machine Learning



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$^{\neg\vee}$ rules

$$p_1(\mathbf{X}_1) \vee \ldots \vee p_n(\mathbf{X}_n) \leftarrow$$
$$r_1(\mathbf{Y}_1), \ldots, r_m(\mathbf{Y}_m), s_1(\mathbf{Z}_1), \ldots, s_k(\mathbf{Z}_k), \neg u_1(\mathbf{W}_1), \ldots, \neg u_h(\mathbf{W}_h)$$

satisfying the following properties

- Datalog safeness: every variable occurring in a rule must appear in at least one of the atoms $r_1(Y_1), \ldots, r_m(Y_m), s_1(Z_1), \ldots, s_k(Z_k)$
- DL weak safeness: every head variable of a rule must appear in at least one of the atoms $r_1(Y_1), \ldots, r_m(Y_m)$

Combining LP & DLs with DL+log: semantics

- FOL-semantics
  - OWA for both DL and Datalog predicates
- NM-semantics: extends stable model semantics of Datalog$^{\neg\vee}$
  - OWA for DL-predicates
  - CWA for Datalog-predicates
- In both semantics, entailment can be reduced to satisfiability
- In Datalog$^{\vee}$, 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$^{\neg\vee}$: A Datalog$^{\neg\vee}$ 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] $\mathrm{RICH} \sqcap \mathrm{UNMARRIED} \sqsubseteq \exists\, \mathrm{WANTS\text{-}TO\text{-}MARRY}^{-}.\top$

[R1] RICH(X) ← famous(X), ¬scientist(X)

$\mathcal{K}$

UNMARRIED(Mary)
UNMARRIED(Joe)

famous(Mary)
famous(Paul)
famous(Joe)
scientist(Joe)

$\mathcal{F}$

$\mathcal{L}^{\text{happy}}$

- {famous/1,RICH/1, WANTS-TO-MARRY/2, LIKES/2}
- happy(X) ← famous(X), WANTS-TO-MARRY(Y,X)

$\mathcal{L}^{\text{LONER}}$

- {famous/1,scientist/1,UNMARRIED/1}
- LONER(X) ← ¬famous(X)

Inducing SHIQ+log rules with ILP: the hypothesis ordering

- SHIQ+log¬ KB $\mathcal{K}$
- SHIQ+log¬ rules H_1, $H_2 \in \mathcal{L}$
- Skolem substitution σ for H_2 w.r.t. $\{H_1\} \cup \mathcal{K}$

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

- $head(H_1)\theta = head(H_2)\sigma$
- $\mathcal{K} \cup body(H_2)\sigma \models body(H_1)\theta$

Generality order boils down to CQ answering!

Inducing SHIQ+log rules with ILP: the hypothesis ordering (2)

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

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

- H_1^{happy} = happy(A) $\leftarrow$ RICH(A)
- H_2^{happy} = happy(X) $\leftarrow$ famous(X)

- $H_1^{happy} \not\succeq_{\mathcal{K}} H_2^{happy}$
- $H_2^{happy} \not\succeq_{\mathcal{K}} H_1^{happy}$

Inducing SHIQ+log rules with ILP: the coverage relations

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

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

Coverage boils down to CQ answering!

Inducing SHIQ+log rules with ILP: the coverage relations (2)

$\mathcal{K}$	$\mathcal{F}_{Mary}$
[A1] $RICH \sqcap UNMARRIED \sqsubseteq \exists$ WANTS-TO-MARRY^{-}.T	UNMARRIED(Mary)
[R1] RICH(X) $\leftarrow$ famous(X), $\neg$ scientist(X)	famous(Mary)

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

covers o_{Mary} = (happy(Mary), $\mathcal{F}_{Mary}$) because

$\mathcal{K} \cup \mathcal{F}_{Mary} \cup H \models$ happy(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

	Learning in Carin-$\mathcal{ALN}$ [24]	Learning in $\mathcal{AL}$-log [15]	Learning in $\mathcal{SHIQ}$+log$^\neg$
prior knowledge	CARIN-$\mathcal{ALN}$ KB	$\mathcal{AL}$-log KB	$\mathcal{SHIQ}$+log$^\neg$ KB
ontology lang.	$\mathcal{ALN}$	$\mathcal{ALC}$	$\mathcal{SHIQ}$
rule lang.	Horn clauses	DATALOG clauses	DATALOG$^\neg$ clauses
hypothesis lang.	CARIN-$\mathcal{ALN}$ non-recursive rules	constrained DATALOG clauses	$\mathcal{SHIQ}$+log$^\neg$ non-recursive rules
target predicate	Horn literal	DATALOG literal	$\mathcal{SHIQ}$/DATALOG literal
observations	interpretations	interpretations/implications	interpretations
induction	predictive	predictive/descriptive	predictive/descriptive
generality order	extension of [3] to CARIN-$\mathcal{ALN}$	extension of [3] to $\mathcal{AL}$-log	extension of [3] to $\mathcal{SHIQ}$+log$^\neg$
coverage test	CARIN-$\mathcal{ALN}$ query answering	$\mathcal{AL}$-log query answering	$\mathcal{SHIQ}$+log$^\neg$ query answering
ref. operators	no	downward	no
implementation	no	partially	no
application	no	yes	no

1. W. Buntine. Generalized subsumption and its application to induction and redundancy. *Artificial Intelligence*, 36(2):149–176, 1988.
2. F.A. Lisi. Building Rules on Top of Ontologies for the Semantic Web with Inductive Logic Programming. *Theory and Practice of Logic Programming*, 8(03):271–300, 2008.
3. C. Rouveirol and V. Ventos. Towards Learning in CARIN-$\mathcal{ALN}$. In J. Cussens and A. Frisch, editors, *Inductive Logic Programming*, volume 1866 of *Lecture Notes in Artificial Intelligence*, pages 191–208. Springer, 2000.

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$^{\neg\vee}$ 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