Logics in Machine Learning and Data Mining: Achievements and Open Issues

Francesca A. Lisi

University of Bari "Aldo Moro"
Department of Computer Science
Lab of Knowledge Acquisition and Machine Learning (LACAM)

francesca.lisi@uniba.it

June 19, 2019

Overview

1. Introduction
2. Three cases for Logics in ML/DM
   - Combining rules and ontologies
   - Dealing with imprecision and granularity
   - Modeling and metamodeling
3. Final remarks
4. References

Introduction I

Model-free AI vs. Model-based AI

- Current hype about AI due to many successful ML applications
- Deep learning follows the model-free approach
  - ML tasks are function-fitting problems!
- Need to construct and use models ("good old-fashioned AI")
  - Logics and probability as main tools
- Is there any ML algorithm following the model-based approach?

Introduction II

Inductive Logic Programming in a nutshell

- Major logic-based approach to rule learning [Muggleton, 1990]
  - Concept Learning within the LP framework
  - Use of background knowledge (BK)
- Bunch of techniques for structuring, searching and bounding the hypothesis space [Nienhuys-Cheng and de Wolf, 1997]
- Two representative ILP algorithms:
  1. FOIL [Quinlan, 1990]
  2. WARMR [Dehaspe and Toivonen, 1999]
- Several extensions, *e.g.*, towards statistical learning and other probabilistic approaches (see [Riguzzi et al., 2014] for a survey).

Three cases for Logics in ML/DM

1. Combining rules and ontologies
2. Dealing with imprecision and granularity
3. Modeling and metamodeling

Combining rules and ontologies I

Logic Programming vs. Descripion Logics

1. CWA vs OWA
2. Single vs. multiple models
3. Negation as failure vs. classical negation
4. Strong negation vs. classical negation
5. Treatment of equality
    - Unique Names Assumption (UNA) [Reiter, 1980] might not hold in DLs
6. Existential quantification
    - Recent work on $\text{D}\scriptstyle{\text{ATALOG}}^{\pm}$ [Calì et al., 2009]
7. Decidability

Combining rules and ontologies II

Exploiting the power of combination

- Combination is more than the sum of the parts
- Very expressive FOL languages as an outcome
- Solutions to the semantic mismatch between LP and DLs

Approaches

Non-hybrid Combination within a *homogeneous* semantic framework

- *e.g.*, description logic programs [Grosof et al., 2003]

Hybrid Combination within a *heterogeneous* semantic framework

- *e.g.*, $\mathcal{AL}$-LOG [Donini et al., 1998]

Combining rules and ontologies III

Learning hybrid rules with ILP

	Learning CARIN-$\mathcal{ALN}$ rules [Rouveirol and Ventos, 2000, Kietz, 2003]	Learning $\mathcal{AL}$-LOG rules [Lisi, 2008]	Learning $\mathcal{SHIQ}$+LOG rules [Lisi, 2010]
prior knowledge	CARIN-$\mathcal{ALN}$ KB	$\mathcal{AL}$-LOG KB	$\mathcal{SHIQ}$+LOG KB
ontology language	$\mathcal{ALN}$	$\mathcal{ALC}$	$\mathcal{SHIQ}$
rule language	HCL	DATALOG	DATALOG
hypothesis language	CARIN-$\mathcal{ALN}$ non-recursive rules	$\mathcal{AL}$-LOG non-recursive rules	$\mathcal{SHIQ}$+LOG non-recursive rules
target predicate	Horn predicate	DATALOG predicate	$\mathcal{SHIQ}$/DATALOG predicate
logical setting	interpretations	interpretations/entailment	entailment
scope of induction	prediction	prediction/description	prediction/description
generality order	generalized subsumption	generalized subsumption	generalized subsumption
coverage test	CARIN query answering	$\mathcal{AL}$-LOG query answering	$\mathcal{DL}$+LOG$^{\vee}$ query answering
ref. operators	n.a.	downward	downward/upward
implementation	unknown	yes, see [Lisi, 2011]	no
application	no	yes, see [Lisi and Malerba, 2004]	no

Combining rules and ontologies IV

$\mathcal{AL}$-QuIn: general features [Lisi, 2011]

- Task: multi-level association rule mining
- Method: levelwise search
- BK: hybrid (relational DB + ontology)
- Knowledge representation formalism: $\mathcal{AL}$-log.
- Upgrade from: Warmr.

Combining rules and ontologies V

$\mathcal{AL}$-QuIn: problem statement

Given

- a relational data set Π
- a taxonomic ontology Σ
- a multi-grained language $\mathcal{L} = \{\mathcal{L}^l\}_{1 \leq l \leq maxG}$
- a set $\{minsup^l\}_{1 \leq l \leq maxG}$ of support thresholds

the problem of *frequent pattern discovery in* Π *at* l *levels of description granularity w.r.t.* Σ, $1 \leq l \leq maxG$, is to find the set $\mathcal{F}$ of all the patterns $P \in \mathcal{L}^l$ frequent in $\mathcal{B} = (\Pi, \Sigma)$, namely P's with support s such that (i) $s \geq minsup^l$ and (ii) all ancestors of P w.r.t. Σ are frequent.

Combining rules and ontologies VI

$\mathcal{AL}$-QuIn: An example

Q_1= q(X) $\leftarrow$ believes(X,Y) & MiddleEastCountry(X)
Q_3= q(X) $\leftarrow$ believes(X,Y) & MiddleEastCountry(X), Religion(Y)
Q_4= q(X) $\leftarrow$ believes(X,Y) & MiddleEastCountry(X), MonotheisticReligion(Y)
Q_5= q(X) $\leftarrow$ believes(X,Y), speaks(X,Z) &
MiddleEastCountry(X), MonotheisticReligion(Y), IndoEuropeanLanguage(Z)

Dealing with imprecision and granularity I

Fuzzy Description Logics

- Several ways of extending DLs with fuzzy logic [Straccia, 2013]
- Some ad-hoc reasoners already available (*e.g.*, *FuzzyDL* [Bobillo and Straccia, 2008])
- Fuzzy quantifiers in Fuzzy DLs [Sanchez and Tettamanzi, 2006]
- Proposal of Fuzzy OWL 2 [Bobillo and Straccia, 2010]

Dealing with imprecision and granularity II

Fuzzy $\mathcal{EL}(\mathbf{D})$ [Straccia, 2005]

- Complex concepts built according to the following syntactic rules:

$$C \rightarrow \top \mid \bot \mid A \mid C_1 \sqcap C_2 \mid \exists R.C \mid \exists T.\mathbf{d}$$

  where

  - $\mathbf{d}$ can be one of the membership functions of fuzzy sets
  - $\sqcap$ and $\exists$ are interpreted as truth combination functions

- Concepts are interpreted as fuzzy sets
- Axioms are graded, *i.e.* have a truth degree α (if omitted, $\alpha = 1$)
  - e.g., $\mathcal{I}$ satisfies an axiom $\langle a{:}C, \alpha\rangle$ if $C^{\mathcal{I}}(a^{\mathcal{I}}) \geq \alpha$
- The *best entailment degree* of an axiom τ w.r.t. $\mathcal{K}$ is defined as

$$bed(\mathcal{K}, \tau) = \sup\{\alpha \mid \mathcal{K} \models \langle \tau, \alpha\rangle\}\ . \tag{1}$$

  For a crisp axiom τ, we also write $\mathcal{K} \models_+ \tau$ iff $bed(\mathcal{K}, \tau) > 0$.

Dealing with imprecision and granularity III

Fuzzy DL Learning

- [Konstantopoulos and Charalambidis, 2010] propose an ad-hoc translation of fuzzy Łukasiewicz $\mathcal{ALC}$ DL constructs into LP in order to apply a conventional ILP method for rule learning.
  - Unsound method
- [Iglesias and Lehmann, 2011] propose to interface CELOE with the *fuzzyDL* reasoner
  - Uncomparable method
- [Lisi and Straccia, 2013] present a method for learning fuzzy $\mathcal{EL}$ GCI axioms from fuzzy DL-Lite KBs.
  - Unimplemented method

Dealing with imprecision and granularity IV

FOIL-$\mathcal{DL}$: general features [Lisi and Straccia, 2014]

- Task: classification rule mining
- Method: sequential covering
- BK: $\mathcal{DL}$ KB
- Hypothesis description language: fuzzy $\mathcal{EL}(\mathbf{D})$.
- Upgrade from: FOIL.

Dealing with imprecision and granularity V

FOIL-$\mathcal{DL}$: problem statement

Given:

- a consistent $\mathcal{DL}$ KB $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ (the *background theory*);
- an atomic concept A_t (the *target concept*);
- a set $\mathcal{E} = \mathcal{E}^+ \cup \mathcal{E}^-$ of crisp $\mathcal{DL}$ concept assertions labelled as either positive or negative examples for H (the *training set*);
- a set $\mathcal{L_H}$ of fuzzy $\mathcal{EL}(\mathbf{D})$ GCIs of the form $C \sqsubseteq A_t$ (the *language of hypotheses*) where C is a complex concept

Find: a set $\mathcal{H} \subset \mathcal{L_H}$ (a *hypothesis*) such that:

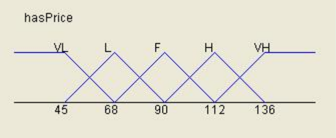
Completeness. $\forall e \in \mathcal{E}^+, \mathcal{K} \cup \mathcal{H} \models_+ e$, and

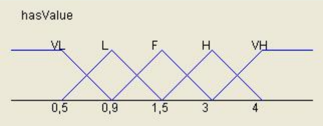
Consistency. $\forall e \in \mathcal{E}^-, \mathcal{K} \cup \mathcal{H} \not\models_+ e$.

Dealing with imprecision and granularity VI

FOIL-$\mathcal{DL}$: An example

Confidence	Axiom
1,000	Hostel subclass of Good_Hotel
1,000	hasPrice_veryhigh subclass of Good_Hotel
0,739	hasDistance some (isDistanceFor some (Bus_Station) and hasValue_low) and hasDistance some (isDistanceFor some (Town_Hall) and hasValue_fair) and hasRank some (Rank) and hasPrice_verylow subclass of Good_Hotel
0,569	hasPrice_high subclass of Good_Hotel
0,289	Hotel_3_Stars and hasDistance some (isDistanceFor some (Train_Station) and hasValue_verylow) and hasPrice_fair subclass of Good_Hotel
0,198	Hotel_4_Stars and hasDistance some (isDistanceFor some (Square) and hasValue_high) and hasRank some (Rank) and hasPrice_fair subclass of Good_Hotel





Dealing with imprecision and granularity VII

Information granulation and DLs [Lisi and Mencar, 2017, Lisi and Mencar, 2018]

Support to **both**

coarser-granularity fuzzy quantified sentences such as

"Many hotels have a low distance from attractions"

and *finer-granularity* fuzzy quantified sentences such as

"Hotel Verdi has a low distance from many attractions"

Modeling and metamodeling I

Model+solver approach in ML/DM

- ML/DM problems as either Constraint Satisfaction Problems or Optimization Problems
- Languages for declarative modeling [De Raedt, 2015]
- Use of efficient solvers, *e.g.*, CP solver in itemset mining [Guns et al., 2011]

Meta-Interpretive Learning

- Novel and promising ILP framework [Muggleton and Lin, 2013]
- Meta-rules (expressed in a higher-order dyadic DATALOG fragment) with procedural constraints
- Meta-interpreter implemented by relying, *e.g.*, on ASP solvers

Modeling and metamodeling II

Higher-order DLs

- [Pan and Horrocks, 2006] propose a stratified Higher-order DL (OWL FA) to cope with meta-assertions about concepts and roles.
- [Motik, 2007] proves that satisfiability in Higher-order $\mathcal{ALCO}$, which is a fragment of OWL Full, is undecidable.
- [De Giacomo et al., 2011] augment a DL with variables that may be interpreted - in a Henkin semantics - as individuals, concepts, and roles at the same time, obtaining a new logic $Hi(\mathcal{DL})$
- [Colucci et al., 2010] introduce second-order features in DLs under Henkin semantics for modeling several non-standard reasoning tasks.

Modeling and metamodeling III

Higher-order DLs in ML/DM

- [Lisi, 2013] extends [Colucci et al., 2010] to some variants of concept learning, thus being the first to propose higher-order DLs as a means for metamodeling in ML.
- [Lisi, 2018] proposes a metaquerying language for mining the Web of Data

Final remarks

Need for integration

1. Learning & Reasoning
2. Symbolic & Sub-symbolic

Angry Birds AI competition

- Simplified and controlled environment for developing future AI systems able to interact with the physical world
- Challenge for AI: Predict the outcome of physical actions without complete knowledge of the world

Open issues in logic-based ML/DM

- scalability
- efficiency

References I

Bobillo, F. and Straccia, U. (2008).
fuzzyDL: An expressive fuzzy description logic reasoner.
In *FUZZ-IEEE 2008, IEEE International Conference on Fuzzy Systems, Hong Kong, China, 1-6 June, 2008, Proceedings*, pages 923–930. IEEE.

Bobillo, F. and Straccia, U. (2010).
Representing fuzzy ontologies in OWL 2.
In *FUZZ-IEEE 2010, IEEE International Conference on Fuzzy Systems, Barcelona, Spain, 18-23 July, 2010, Proceedings*, pages 1–6. IEEE.

Calì, A., Gottlob, G., and Lukasiewicz, T. (2009).
Datalog$^{\pm}$: a unified approach to ontologies and integrity constraints.
In Fagin, R., editor, *Database Theory - ICDT 2009, 12th International Conference, St. Petersburg, Russia, March 23-25, 2009, Proceedings*, volume 361 of *ACM International Conference Proceeding Series*, pages 14–30. ACM.

Colucci, S., Di Noia, T., Di Sciascio, E., Donini, F. M., and Ragone, A. (2010).
A unified framework for non-standard reasoning services in description logics.
In Coelho, H., Studer, R., and Wooldridge, M., editors, *ECAI 2010 - 19th European Conference on Artificial Intelligence, Lisbon, Portugal, August 16-20, 2010, Proceedings*, volume 215 of *Frontiers in Artificial Intelligence and Applications*, pages 479–484. IOS Press.

De Giacomo, G., Lenzerini, M., and Rosati, R. (2011).
Higher-order description logics for domain metamodeling.
In Burgard, W. and Roth, D., editors, *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2011, San Francisco, California, USA, August 7-11, 2011*.

De Raedt, L. (2015).
Languages for learning and mining.
In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, pages 4107–4111.

Dehaspe, L. and Toivonen, H. (1999).
Discovery of frequent DATALOG patterns.
Data Mining and Knowledge Discovery, 3:7–36.

Donini, F. M., Lenzerini, M., Nardi, D., and Schaerf, A. (1998).
$\mathcal{AL}$-log: Integrating Datalog and Description Logics.
Journal of Intelligent Information Systems, 10(3):227–252.

Grosof, B. N., Horrocks, I., Volz, R., and Decker, S. (2003).
Description logic programs: combining logic programs with description logic.
In *Proceedings of the 12th International World Wide Web Conference*, pages 48–57. ACM.

Guns, T., Nijssen, S., and De Raedt, L. (2011).
Itemset mining: A constraint programming perspective.
Artificial Intelligence, 175(12-13):1951–1983.

Iglesias, J. and Lehmann, J. (2011).
Towards integrating fuzzy logic capabilities into an ontology-based inductive logic programming framework.
In *Proc. of the 11th Int. Conf. on Intelligent Systems Design and Applications*. IEEE Press.

Kietz, J.-U. (2003).
Learnability of description logic programs.
In Matwin, S. and Sammut, C., editors, *Inductive Logic Programming, 12th International Conference, ILP 2002, Sydney, Australia, July 9-11, 2002. Revised Papers*, volume 2583 of *Lecture Notes in Computer Science*, pages 117–132. Springer.

References III

Konstantopoulos, S. and Charalambidis, A. (2010).
Formulating description logic learning as an inductive logic programming task.
In *Proc. of the 19th IEEE Int. Conf. on Fuzzy Systems*, pages 1–7. IEEE Press.

Lisi, F. A. (2008).
Building Rules on Top of Ontologies for the Semantic Web with Inductive Logic Programming.
Theory and Practice of Logic Programming, 8(03):271–300.

Lisi, F. A. (2010).
Inductive Logic Programming in Databases: From Datalog to $\mathcal{DL}$+log.
Theory and Practice of Logic Programming, 10(3):331–359.

Lisi, F. A. (2011).
$\mathcal{AL}$-QuIn: An Onto-Relational Learning System for Semantic Web Mining.
International Journal on Semantic Web and Information Systems, 7(3):1–22.

Lisi, F. A. (2013).
A declarative modeling language for concept learning in description logics.
In Riguzzi, F. and Zelezny, F., editors, *Inductive Logic Programming, 22nd International Conference, ILP 2012, Dubrovnik, Croatia, September 17-19, 2012, Revised Selected Papers*, volume 7842 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg.

Lisi, F. A. (2018).
Mining the web of data with metaqueries.
In Riguzzi, F., Bellodi, E., and Zese, R., editors, *Up-and-Coming and Short Papers of the 28th International Conference on Inductive Logic Programming (ILP 2018), Ferrara, Italy, September 2-4, 2018*, volume 2206 of *CEUR Workshop Proceedings*, pages 92–99. CEUR-WS.org.

References IV

Lisi, F. A. and Malerba, D. (2004).
Inducing Multi-Level Association Rules from Multiple Relations.
Machine Learning, 55:175–210.

Lisi, F. A. and Mencar, C. (2017).
Introducing fuzzy quantification in OWL 2 ontologies.
In Monica, D. D., Murano, A., Rubin, S., and Sauro, L., editors, *Joint Proceedings of the 18th Italian Conference on Theoretical Computer Science and the 32nd Italian Conference on Computational Logic co-located with the 2017 IEEE International Workshop on Measurements and Networking (2017 IEEE M&N), Naples, Italy, September 26-28, 2017.*, volume 1949 of *CEUR Workshop Proceedings*, pages 321–325. CEUR-WS.org.

Lisi, F. A. and Mencar, C. (2018).
A granular computing method for OWL ontologies.
Fundamenta Informaticae, 159(1–2):147–174.

Lisi, F. A. and Straccia, U. (2013).
A logic-based computational method for the automated induction of fuzzy ontology axioms.
Fundamenta Informaticae, 124(4):503–519.

Lisi, F. A. and Straccia, U. (2014).
A FOIL-like Method for Learning under Incompleteness and Vagueness.
In Zaverucha, G., Santos Costa, V., and Paes, A., editors, *Inductive Logic Programming - 23rd International Conference, ILP 2013, Rio de Janeiro, Brazil, August 28-30, 2013, Revised Selected Papers*, volume 8812 of *Lecture Notes in Computer Science*, pages 123–139. Springer.

Motik, B. (2007).
On the properties of metamodeling in OWL.
Journal of Logic and Computation, 17(4):617–637.

References V

Muggleton, S. H. (1990).
Inductive logic programming.
In Arikawa, S., Goto, S., Ohsuga, S., and Yokomori, T., editors, *Proceedings of the 1st Conference on Algorithmic Learning Theory*. Springer/Ohmsma.

Muggleton, S. H. and Lin, D. (2013).
Meta-interpretive learning of higher-order dyadic datalog: Predicate invention revisited.
In Rossi, F., editor, *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*. IJCAI/AAAI.

Nienhuys-Cheng, S.-H. and de Wolf, R. (1997).
Foundations of Inductive Logic Programming, volume 1228 of *Lecture Notes in Artificial Intelligence*.
Springer.

Pan, J. Z. and Horrocks, I. (2006).
OWL FA: a metamodeling extension of OWL DL.
In Carr, L., De Roure, D., Iyengar, A., Goble, C. A., and Dahlin, M., editors, *Proceedings of the 15th international conference on World Wide Web, WWW 2006, Edinburgh, Scotland, UK, May 23-26, 2006*, pages 1065–1066. ACM.

Quinlan, J. R. (1990).
Learning logical definitions from relations.
Machine Learning, 5:239–266.

Reiter, R. (1980).
Equality and domain closure in first order databases.
Journal of ACM, 27:235–249.

References VI

Riguzzi, F., Bellodi, E., and Zese, R. (2014).
A history of probabilistic inductive logic programming.
Frontiers in Robotics and AI, 2014.

Rouveirol, C. and Ventos, V. (2000).
Towards Learning in CARIN-$\mathcal{ALN}$.
In Cussens, J. and Frisch, A. M., editors, *Inductive Logic Programming, 10th International Conference, ILP 2000, London, UK, July 24-27, 2000, Proceedings*, volume 1866 of *Lecture Notes in Artificial Intelligence*, pages 191–208. Springer.

Sanchez, D. and Tettamanzi, A. G. (2006).
Fuzzy quantification in fuzzy description logics.
In Sanchez, E., editor, *Fuzzy Logic and the Semantic Web*, volume 1 of *Capturing Intelligence*, pages 135 – 159. Elsevier.

Straccia, U. (2005).
Description logics with fuzzy concrete domains.
In *UAI '05, Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, Edinburgh, Scotland, July 26-29, 2005*, pages 559–567. AUAI Press.

Straccia, U. (2013).
Foundations of Fuzzy Logic and Semantic Web Languages.
CRC Studies in Informatics Series. Chapman & Hall.