

# Answering Queries in Description Logics: Theory and Applications to Data Management

Diego Calvanese<sup>1</sup>, Michael Zakharyashev<sup>2</sup>

<sup>1</sup> KRDB Research Centre  
Free University of Bozen-Bolzano

<sup>2</sup> Birbeck College, London

ESLLI 2010, August 14–20, 2010  
Copenhagen, Denmark

# Overview of the Course

- 1 Introduction and background
  - 1 Ontology-based data management
  - 2 Brief introduction to computational complexity
  - 3 Query answering in databases
  - 4 Querying databases and ontologies
- 2 Lightweight description logics
  - 5 Introduction to description logics
  - 6 DLs for conceptual data modeling: the *DL-Lite* family
  - 7 The  $\mathcal{EL}$  family of tractable description logics
- 3 Query answering in the *DL-Lite* family
  - 8 Query answering in description logics
  - 9 Lower bounds for description logics beyond *DL-Lite*
  - 10 Reasoning and query answering by rewriting
- 4 The combined approach to query answering
  - 11 Query answering in *DL-Lite*: data completion
  - 12 Query rewriting in  $\mathcal{EL}$
- 5 Linking ontologies to relational data
  - 13 The impedance mismatch problem
  - 14 Query answering in Ontology-Based Data Access systems
- 6 Conclusions and references



## Conclusions and references

# Main publications

The results presented in this course have been published in the following papers:

- Reasoning and query answering in *DL-Lite* and relatives: [Artale *et al.*, 2009; Calvanese *et al.*, 2007a; Calvanese *et al.*, 2009a]
- Reasoning and query answering in  $\mathcal{EL}$  and extensions: [Baader *et al.*, 2005; Baader *et al.*, 2008; Lutz *et al.*, 2010; Lutz *et al.*, 2009]
- Query answering in *DL-Lite* using data completion: [Kontchakov *et al.*, 2010; Rosati and Almatelli, 2010]
- Mapping to data sources and OBDA: [Poggi *et al.*, 2008a]

Additional related material:

- Connection between description logics and conceptual modeling formalisms: [Calvanese *et al.*, 1998; Berardi *et al.*, 2005; Artale *et al.*, 2007; Calvanese *et al.*, 2009b]
- Descriptions of the QUONTO/MASTRO Tool : [Acciarri *et al.*, 2005; Poggi *et al.*, 2008b; Rodríguez-Muro and Calvanese, 2008]



## Further theoretical work

The results presented in this course have also inspired additional work relevant for ontology-based data access:

- Alternative query rewriting techniques based on resolution for more expressive logics (with recursive rewritings) [Pérez-Urbina *et al.*, 2009].
- Query rewriting techniques for database inspired constraint languages [Calì *et al.*, 2009a; Calì *et al.*, 2009b].
- We have considered mainly query answering. However, several other ontology-based services are of importance:
  - write-also access: updating a data source through an ontology [De Giacomo *et al.*, 2009; Calvanese *et al.*, 2010]
  - modularity and minimal module extraction [Kontchakov *et al.*, 2008; Kontchakov *et al.*, 2009]
  - provenance and explanation [Borgida *et al.*, 2008]
- Reasoning with respect to finite models only [Rosati, 2008].
- We have dealt only with the static aspects of information systems. However a crucial issue is how to deal with **dynamic aspects**. Preliminary results are in [Calvanese *et al.*, 2007b]. The general problem is largely unexplored.

Work on most of these issues is still ongoing.



## Further practical and experimental work

The theoretical results indicate a good computational behaviour in the size of the data. However, performance is a critical issue in practice:

- The rewriting consists of a large number of CQs. Query containment can be used to prune the rewriting. This is already implemented in the QUONTO system, but requires further optimizations.
- The SQL queries generated by the mapping unfolding are not easy to process by the DBMS engine (e.g., they may contain complex joins on skolem terms computed on the fly).  
Different mapping unfolding strategies have a strong impact on computational complexity. Experimentation is ongoing to assess the tradeoff.
- Further extensive experimentations are ongoing:
  - on artificially generated data;
  - on real-world use cases.

# Acknowledgements

- Roman Kontchakov (Birbeck College)
- Carsten Lutz (Univ. of Bremen)
- Frank Wolter (Univ. of Liverpool)
- Sapienza Università di Roma
  - Claudio Corona
  - Giuseppe De Giacomo
  - Domenico Lembo
  - Maurizio Lenzerini
  - Antonella Poggi
  - Riccardo Rosati
  - Marco Ruzzi
  - Domenico Fabio Savo
- Free University of Bozen-Bolzano
  - Alessandro Artale
  - Mariano Rodriguez-Muro
  - Manfred Gerstgrasser
- Many students (thanks!)



# References I

[Acciarri *et al.*, 2005] Andrea Acciarri, Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Mattia Palmieri, and Riccardo Rosati.

QUONTO: Querying ONTOlogies.

In *Proc. of the 20th Nat. Conf. on Artificial Intelligence (AAAI 2005)*, pages 1670–1671, 2005.

[Artale *et al.*, 2007] Alessandro Artale, Diego Calvanese, Roman Kontchakov, Vladislav Ryzhikov, and Michael Zakharyashev.

Reasoning over extended ER models.

In *Proc. of the 26th Int. Conf. on Conceptual Modeling (ER 2007)*, volume 4801 of *Lecture Notes in Computer Science*, pages 277–292. Springer, 2007.

[Artale *et al.*, 2009] Alessandro Artale, Diego Calvanese, Roman Kontchakov, and Michael Zakharyashev.

The *DL-Lite* family and relations.

*J. of Artificial Intelligence Research*, 36:1–69, 2009.



# References II

- [Baader *et al.*, 2005] Franz Baader, Sebastian Brandt, and Carsten Lutz.  
Pushing the  $\mathcal{EL}$  envelope.  
In *Proc. of the 19th Int. Joint Conf. on Artificial Intelligence (IJCAI 2005)*, pages 364–369, 2005.
- [Baader *et al.*, 2008] Franz Baader, Sebastian Brandt, and Carsten Lutz.  
Pushing the  $\mathcal{EL}$  envelope further.  
In Kendall Clark and Peter F. Patel-Schneider, editors, *Proc. of the 4th Int. Workshop on OWL: Experiences and Directions (OWLED 2008 DC)*, 2008.
- [Berardi *et al.*, 2005] Daniela Berardi, Diego Calvanese, and Giuseppe De Giacomo.  
Reasoning on UML class diagrams.  
*Artificial Intelligence*, 168(1–2):70–118, 2005.
- [Borgida *et al.*, 2008] Alexander Borgida, Diego Calvanese, and Mariano Rodríguez-Muro.  
Explanation in the *DL-Lite* family of description logics.  
In *Proc. of the 7th Int. Conf. on Ontologies, DataBases, and Applications of Semantics (ODBASE 2008)*, volume 5332 of *Lecture Notes in Computer Science*, pages 1440–1457. Springer, 2008.

## References III

- [Calì *et al.*, 2009a] Andrea Calì, Georg Gottlob, and Thomas Lukasiewicz.  
Datalog<sup>±</sup>: a unified approach to ontologies and integrity constraints.  
In *Proc. of the 12th Int. Conf. on Database Theory (ICDT 2009)*, pages 14–30, 2009.
- [Calì *et al.*, 2009b] Andrea Calì, Georg Gottlob, and Thomas Lukasiewicz.  
A general Datalog-based framework for tractable query answering over ontologies.  
In *Proc. of the 28th ACM SIGACT SIGMOD SIGART Symp. on Principles of Database Systems (PODS 2009)*, pages 77–86, 2009.
- [Calvanese *et al.*, 1998] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi.  
Description logics for conceptual data modeling.  
In Jan Chomicki and Günter Saake, editors, *Logics for Databases and Information Systems*, pages 229–264. Kluwer Academic Publishers, 1998.
- [Calvanese *et al.*, 2007a] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati.  
Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family.  
*J. of Automated Reasoning*, 39(3):385–429, 2007.

## References IV

- [Calvanese *et al.*, 2007b] Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Riccardo Rosati.

Actions and programs over description logic ontologies.

In *Proc. of the 20th Int. Workshop on Description Logic (DL 2007)*, volume 250 of *CEUR Electronic Workshop Proceedings*, <http://ceur-ws.org/>, pages 29–40, 2007.

- [Calvanese *et al.*, 2009a] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Mariano Rodríguez-Muro, and Riccardo Rosati.

Ontologies and databases: The *DL-Lite* approach.

In Sergio Tessaris and Enrico Franconi, editors, *Semantic Technologies for Informations Systems – 5th Int. Reasoning Web Summer School (RW 2009)*, volume 5689 of *Lecture Notes in Computer Science*, pages 255–356. Springer, 2009.

- [Calvanese *et al.*, 2009b] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati.

Conceptual modeling for data integration.

In Alex T. Borgida, Vinay Chaudhri, Paolo Giorgini, and Eric Yu, editors, *Conceptual Modeling: Foundations and Applications – Essays in Honor of John Mylopoulos*, volume 5600 of *Lecture Notes in Computer Science*, pages 173–197. Springer, 2009.

# References V

- [Calvanese *et al.*, 2010] Diego Calvanese, Evgeny Kharlamov, Werner Nutt, and Dmitriy Zheleznyakov.  
Evolution of *DL-Lite* knowledge bases.  
In *Proc. of the 9th Int. Semantic Web Conf. (ISWC 2010)*, 2010.
- [De Giacomo *et al.*, 2009] Giuseppe De Giacomo, Maurizio Lenzerini, Antonella Poggi, and Riccardo Rosati.  
On instance-level update and erasure in description logic ontologies.  
*J. of Logic and Computation, Special Issue on Ontology Dynamics*, 19(5):745–770, 2009.
- [Kontchakov *et al.*, 2008] Roman Kontchakov, Frank Wolter, and Michael Zakharyashev.  
Can you tell the difference between *DL-Lite* ontologies?  
In *Proc. of the 11th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2008)*, pages 285–295, 2008.
- [Kontchakov *et al.*, 2009] R. Kontchakov, L. Pulina, U. Sattler, T. Schneider, P. Selmer, F. Wolter, and M. Zakharyashev.  
Minimal module extraction from *DL-Lite* ontologies using QBF solvers.  
In *Proc. of the 21st Int. Joint Conf. on Artificial Intelligence (IJCAI 2009)*, pages 836–840, 2009.

# References VI

[Kontchakov *et al.*, 2010] Roman Kontchakov, Carsten Lutz, David Toman, Frank Wolter, and Michael Zakharyashev.

The combined approach to query answering in *DL-Lite*.

In *Proc. of the 12th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2010)*, 2010.

[Lutz *et al.*, 2009] Carsten Lutz, David Toman, and Frank Wolter.

Conjunctive query answering in the description logic  $\mathcal{EL}$  using a relational database system.

In *Proc. of the 21st Int. Joint Conf. on Artificial Intelligence (IJCAI 2009)*, pages 2070–2075, 2009.

[Lutz *et al.*, 2010] Carsten Lutz, Robert Piro, and Frank Wolter.

Enriching  $\mathcal{EL}$ -concepts with greatest fixpoints.

In *Proc. of the 19th Eur. Conf. on Artificial Intelligence (ECAI 2010)*, 2010.

[Pérez-Urbina *et al.*, 2009] Héctor Pérez-Urbina, Boris Motik, and Ian Horrocks.

A comparison of query rewriting techniques for *DL-lite*.

In *Proc. of the 22nd Int. Workshop on Description Logic (DL 2009)*, volume 477 of *CEUR Electronic Workshop Proceedings*, <http://ceur-ws.org/>, 2009.

## References VII

[Poggi *et al.*, 2008a] Antonella Poggi, Domenico Lembo, Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Riccardo Rosati.

Linking data to ontologies.

*J. on Data Semantics*, X:133–173, 2008.

[Poggi *et al.*, 2008b] Antonella Poggi, Mariano Rodríguez-Muro, and Marco Ruzzi.

Ontology-based database access with DIG-Mastro and the OBDA Plugin for Protégé.

In Kendall Clark and Peter F. Patel-Schneider, editors, *Proc. of the 4th Int. Workshop on OWL: Experiences and Directions (OWLED 2008 DC)*, 2008.

[Rodríguez-Muro and Calvanese, 2008] Mariano Rodríguez-Muro and Diego Calvanese.

Towards an open framework for ontology based data access with Protégé and DIG 1.1.

In *Proc. of the 5th Int. Workshop on OWL: Experiences and Directions (OWLED 2008)*, volume 432 of *CEUR Electronic Workshop Proceedings*, <http://ceur-ws.org/>, 2008.

[Rosati and Almatelli, 2010] Riccardo Rosati and Alessandro Almatelli.

Improving query answering over *DL-Lite* ontologies.

In *Proc. of the 12th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2010)*, 2010.

## References VIII

[Rosati, 2008] Ricardo Rosati.

Finite model reasoning in *DL-Lite*.

In *Proc. of the 5th European Semantic Web Conf. (ESWC 2008)*, 2008.