

Practical Inconsistency Handling in Expressive Settings

Postdoc Position in Bordeaux, France

Keywords

Inconsistency handling, data quality, ontology-based data access, query answering, description logics, knowledge representation and reasoning

Context

Accessing the relevant information in real-world data is difficult, time-consuming, and error-prone due to the need to integrate data across multiple heterogeneous sources. Moreover, even if this first hurdle is overcome, a perhaps even more daunting challenge arises: how to obtain reliable insights from imperfect data? It is widely acknowledged that real-world data is plagued with quality issues, such as incompleteness (missing information) and errors (false or outdated information).

The ontology-based data access (OBDA) paradigm [1,2,3] addresses the first challenge by facilitating access to (potentially heterogeneous) data sources through the use of ontologies that specify a convenient user-friendly vocabulary for query formulation (which abstracts from the way the data is stored) and capture domain knowledge that can be exploited at query time, via automated reasoning, to obtain more complete query results. While OBDA systems are growing in maturity [4], they too often fail to address the data quality issue, aside from issuing warnings when inconsistencies are discovered.

To widen the applicability of OBDA, it is essential to equip OBDA systems with appropriate mechanisms for handling imperfect data: how to obtain meaningful answers to queries posed over imperfect data, and how best to generate a high-quality version of the data? While these questions have begun to be explored [5,6], building upon a long line of work on database repairs [7], we are still far from having robust and widely applicable techniques for handling data quality in OBDA.

The postdoc position is part of the **INTENDED** Chair on Artificial Intelligence, whose aim is to develop intelligent, knowledge-based methods for handling imperfect data. The chair started in September 2020 and has a duration of four years. It is funded by the French National Research Agency (ANR) and the University of Bordeaux.

Position

This is a two-year full-time postdoctoral researcher position, with an approximate monthly net salary of 2100€. The position does not include any mandatory teaching or supervisory activities, but there will be opportunities to engage in teaching and/or student supervision if desired. The position will remain open until a suitable candidate is found.

Research Environment

The position will be based at **LaBRI** (Laboratoire Bordelais de Recherche en Informatique), a computer science research lab located on the University of Bordeaux's Talence campus. The lab can be easily reached from the city center of Bordeaux by tram.

The postdoctoral researcher will participate in the new research team **RATIO** (Reasoning with data, knowledge and constraints), which is part of a much larger Formal Methods & Models department that brings together dozens of researchers interested in applying logical and automata-theoretic methods to a range of problems in computer science.

The postdoctoral researcher will collaborate with Meghyn Bienvenu (leader of the chair) and will have many opportunities to interact with the other INTENDED participants and external collaborators. The chair provides significant resources to finance trips to conferences and to host visiting researchers.

Research Topic

Currently, state-of-the-art inconsistency-tolerant algorithms rely, directly or indirectly, on the first-order rewritability property, which restricts their applicability to simple ontology languages (e.g. DL-Lite, OWL 2 QL). Tackling other ontology languages (such as OWL 2 EL, a popular choice for biomedical ontologies) presents significant technical challenges, as the facts involved in contradictions cannot be efficiently identified.

The postdoctoral researcher will be involved in developing pragmatic approaches to handling inconsistencies in more expressive OBDA settings. This could include conceiving new inconsistency-tolerant query answering algorithms for more expressive ontology languages, devising optimizations, identifying interesting tractable subcases, and extending the algorithms to tackle other important features (e.g. mappings, temporal dimension) that have been little explored thus far.

Depending on the interests and aptitude of the postdoctoral researcher, the work can be entirely foundational, or it could involve a more practical component with the implementation and testing of the developed algorithms.

Note: The precise research topic may be adjusted to suit the background and interests of the postdoctoral researcher, so brilliant candidates with an interest in developing principled approaches to tackling data quality should not hesitate to apply.

Candidate Profile

Applicants should hold (or be close to completing) a PhD degree in computer science and should have an excellent research record in knowledge representation and reasoning or databases (ideally, demonstrated by publications in top-tier venues).

Familiarity with logic and/or ontology languages is desirable.

Experience and interest in one or more of the following areas would be relevant and welcome: description logics, Semantic Web, automated reasoning, inconsistency handling, data integration, data quality, theoretical computer science (complexity).

Strong English language skills (reading, writing, & speaking) are expected, but knowledge of French is not required. The working language can be either French or English.

Contact

Interested candidates should contact Meghyn Bienvenu (meghyn.bienvenu@labri.fr) and provide a detailed CV (with publication list and names of 2-3 references) and a brief description of how the project topic relates to their prior experience and research interests.

Bibliography

[1] A. Poggi, D. Lembo, D. Calvanese, G. De Giacomo, M. Lenzerini, and R. Rosati. Linking data to ontologies. *Journal of Data Semantics*, 10:133–173, 2008.

[2] G. Xiao, D. Calvanese, R. Kontchakov, D. Lembo, A. Poggi, R. Rosati, and M. Zakharyashev. Ontology-based data access: A survey. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 5511–5519, 2018.

[3] M. Bienvenu, M. Ortiz: Ontology-Mediated Query Answering with Data-Tractable Description Logics. *Tutorial Notes of the 11th International Reasoning Web Summer School (LNCS 9203)*, pages 218-307, 2015.

[4] D. Calvanese, B. Cogrel, S. Komla-Ebri, R. Kontchakov, D. Lanti, M. Rezk, M. Rodriguez-Muro, G. Xiao, "Ontop: Answering SPARQL queries over relational databases," *Semantic Web Journal*, vol. 8, no. 3, pp. 471-487, 2017.

[5] M. Bienvenu, C. Bourgaux: Inconsistency-Tolerant Querying of Description Logic Knowledge Bases. *Tutorial Notes of the 12th International Reasoning Web Summer School (LNCS 9885)*, pages 156- 202, 2016.

[6] F. Parisi, A. Pugliese, G. Simari, V. S. Subrahmanian and M. V. Martinez, "Policy-based inconsistency management in relational databases," *International Journal of Approximate Reasoning*, vol. 55, no. 2, pp. 501-528, 2014.

[7] M. Arenas, L. Bertossi, J. Chomicki: Consistent Query Answers in Inconsistent Databases. In *Proceedings of the 18th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS)*: 68-79, 1999.