CRIStAL
Centre de Recherche en Informatique,
Signal et Automatique de Lille

Villeneuve d'Ascq, France

Novembre 29th 2025

Senior Researcher dr. hab Pierre Bourhis

UMR CRIStAL
Université de Lille - Campus scientifique
Bâtiment ESPRIT
Avenue Henri Poincaré
59655 Villeneuve d'Ascq

Review

For the doctoral thesis

Regular Path Queries and Modal Constraints:

Decidability of Query Containment under a Graph Database Schema

By Albert Gutowski

This is a report for the doctoral thesis of Mr Albert Gutowski submitted to the University of Warsaw.

This thesis has a main and major contribution on logical reasoning. The question investigated in this thesis is the problem of containment of union of conjunctive regular path queries under constraints expressed by the ontology language ALCO regular. The problem of containment (recursive queries) is a classical problem in database theory and in knowledge representation and it has been studied for various query languages and different constraints languages. The problem of containment of a query Q1 into another query Q2 under the set of constraints C is defined as follows: for every instance satisfying C if this instance satisfies Q1 then it satisfies Q2. Classically, this problem is studied when instances can be infinite and numerous results have been developed under this assumption. However, this another variant of this problem when considered instances are only finite. Surprisingly, this problem is much more complex and most of the results are related to the property of finite controllability which means that both problems for finite and infinite instances are equivalent. Unfortunately, it may be not be the case when queries are recursive under none trivial constraints.

There are very few existing results for recursive queries and constraints and in general, their proofs are extremely complex.

This thesis studies this problem of containment under constraints for finite instance for a well-accepted query language called Union of conjunctive regular path queries allowing to express combination of path properties between nodes of a graph. The main constraint language called ALCO regular describes constraints on nodes of the graphs. The major part of this thesis is the proof of this complex and impressive result. The manuscript is composed of twelve chapters, the first two introduce the problem and the main notions and the last one concludes the manuscript. The other chapters can be split in two main parts: a first one that introduce the main constructions and tools and use them to proof the problem of union of conjunctive queries under ALCO ontologies and a second part to extend the tools presented in the first part in order to prove the main result of the PhD. Even though the problem of containment of conjunctive queries under ALCO for finite instance can be proven by reducing it to the satisfiability of a formula expressed in UNFO which has the finite controllability property, presenting the concepts in this much simpler but still complex setting is a very elegant and very pedagogical approach. The proof is composed of four different reductions giving the chapters from 3 to 6:

Chapter 3. The reduction from the problem of conjunctive query containment through ALCO constraints to the problem of the homomorphism entailment problem. This chapter presents the definitions and the reduction from one problem to another.

Chapter 4. The reduction from the homomorphism entailment problem to the homomorphism coverage problem for G. The problem of homomorphism coverage is the following: given a graph H and class of graphs C and modal profile S, is there a graph G from C and S such that there is no homomorphism from H to G.

Chapter 5. The reduction from the problem homomorphism coverage problem to the minimal downsets problem. The downset for a class of graphs C and modal profile S is the set of all graphs with at most n edges that map homomorphically to a graph of C and S. The coverage homomorphism problem is true iff the graphs in the minimal downset contains a graph which H has a homomorphism to.

Chapter 6. The reduction from the problem of the minimal downsets problem to the satisfiability problem and the resolution of the satisfiability problem. The satisfiability problem is in this case the problem of knowing if there a graph in the class of C satisfying S.

The second part of the PhD describes the proof for the main result and it is written following the same structure in a mirror approach for more complex structures to consider not only local constraints but the one about reachable nodes from regular queries and constraints. The main intuitive idea is to consider no the classical graph but a graph which is augmented by annotated edges that describes non local relations and the different problems are solved over these structures.

Chapter 7. This chapter introduces a simple augmented annotated graph to consider the reachability between nodes. It shows the reduction of computing minimal downsets of this annotated graph to the satisfiability problem to another augmented annotated graph. The chapter describes well this reduction and the concepts despite its complexity.

Chapter 8. The reduction from the problem of the minimal downsets problem for ranked graphs to the satisfiability problem and the resolution of the satisfiability problem to another augmented graph. From my point of view, this chapter is the most complex one of the PhD as it introduces a new structure, ranked graphs, which are not motivated but they will be used later in the proof. The proofs are also quite complex.

Chapter 9. The reduction of the minimal downsets problem for automata annotated graphs to the same problem for ranked graphs. Automata annotated graphs are a generalization of reachability annotation for automata runs over path of the graphs.

Chapter 10. The reduction of the homomorphism entailment problem for automata annotated to the homomorphism coverage problem for the same kind of annotated graphs.

Chapter 11. The reduction of the problem of UCRPQ containment under constraints expressed by ALCO regular to the homomorphism entailment problem for automata annotated graphs.

....

The approach and the tools developed in this PhD are very interesting and highly complex and for the main parts completely new from my knowledge. I think that this method could be applied to other related problems and could lead to adeeper understanding on how reasoning on finite instances.

The paper is in general well written and very pedagogical and most of the notions and constructions are very well illustrated with examples.

A part of the main result was published in the International Conference on Principles of Knowledge Representation and Reasoning 2022 and won the best Student Award Paper.

I have several comments and questions to the author that are here to show my enthusiasm and my high opinion on this thesis and they might be worth discussing during the defense. Here a selection of them:

- The instances studied in this paper are graphs. Could this approach be lifted to general instances, i.e. hypergraph when the constraints are expressed about the hypergraph: inclusion dependencies, linear dependencies, (frontier)-guarded TGDs?
- Could you explain which are the key points of your approach that do not allow to resolve finite satisfiability of Jung *et al.* (2018)?
- There exist recursive query languages more expressive than CRPQs such as CQPDL denote the language of conjunctive queries where binary can be PDL programs or MQ (Pierre Bourhis, Markus Krötzsch, Sebastian Rudolph: Reasonable Highly Expressive Query Languages IJCAI 2015), which are the limits on your approach to consider these languages?

To conclude, the work presented in this manuscript is very interesting and satisfies without a doubt the common standard for a doctoral thesis and even surpassed them.

Pierre Bourhis