

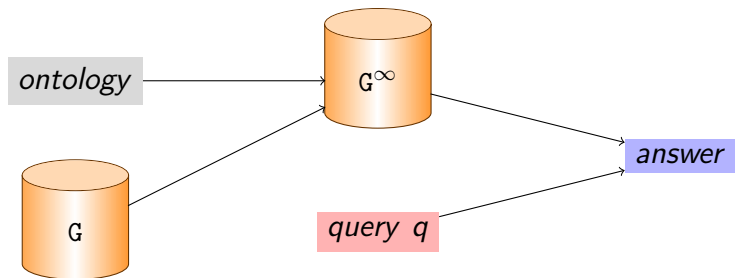
Grouping Answers in Ontology-Based Query Answering: the RDFS Case

Maxime Buron and Michaël Thomazo

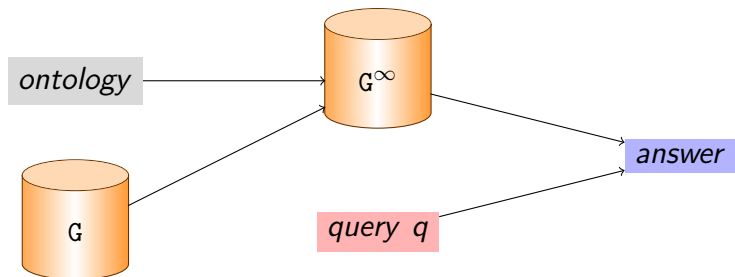
Inria, Université Paris-Saclay
LIX, École Polytechnique, Université Paris-Saclay

WebClaimExplain – June 2nd, 2017

Materialization-based query answering

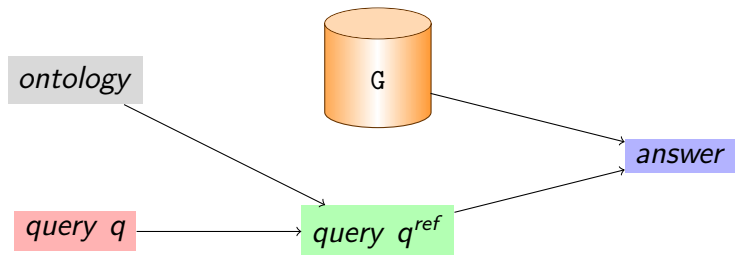


Materialization-based query answering

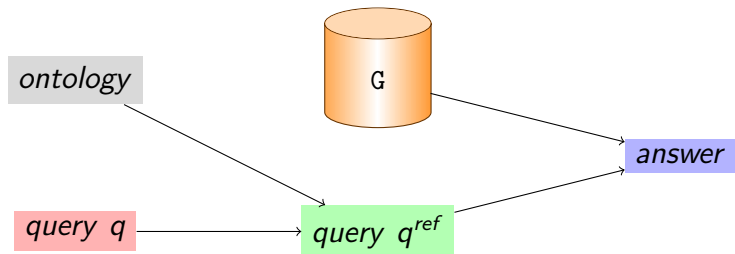


- ▶ $q(G^\infty)$ can be computed using an RDBMS
- ▶ G^∞ needs time to be computed and space to be stored
- ▶ Not suitable for high update rate (data and/or schema triples)

Reformulation-based query answering



Reformulation-based query answering



- ▶ $q^{ref}(G)$ can be evaluated using an RDBMS
- ▶ Robust to updates
- ▶ Reformulated queries are complex, thus costly to evaluate

Ontology Mediated Query Answering

- ▶ **Data:** Professor(Alice), Reviewer(Alice)
- ▶ **Query :** $\exists x \exists y \text{ Teacher}(x) \wedge \text{reviews}(x, y)$

Ontology Mediated Query Answering

- ▶ **Data:** Professor(Alice), Reviewer(Alice)
- ▶ **Query :** $\exists x \exists y \text{ Teacher}(x) \wedge \text{reviews}(x, y)$
- ▶ **Ontology (semantics) :**
 - ▶ $\forall x \text{ Reviewer}(x) \rightarrow \exists y \text{ reviews}(x, y)$
 - ▶ $\forall x \text{ Professor}(x) \rightarrow \text{Teacher}(x)$

Ontology Mediated Query Answering

- ▶ **Data:** Professor(Alice), Reviewer(Alice)
- ▶ **Query :** $\exists x \exists y \text{ Teacher}(x) \wedge \text{reviews}(x, y)$
- ▶ **Ontology (semantics) :**
 - ▶ $\forall x \text{ Reviewer}(x) \rightarrow \exists y \text{ reviews}(x, y)$
 - ▶ $\forall x \text{ Professor}(x) \rightarrow \text{Teacher}(x)$

Materialization

Professor(Alice)

Reviewer(Alice)

Teacher(Alice)

$\exists y_1 \text{ reviews}(Alice, y_1)$

Ontology Mediated Query Answering

- ▶ **Data:** Professor(Alice), Reviewer(Alice)
- ▶ **Query :** $\exists x \exists y \text{ Teacher}(x) \wedge \text{reviews}(x, y)$
- ▶ **Ontology (semantics) :**
 - ▶ $\forall x \text{ Reviewer}(x) \rightarrow \exists y \text{ reviews}(x, y)$
 - ▶ $\forall x \text{ Professor}(x) \rightarrow \text{Teacher}(x)$

Materialization

Professor(Alice)
Reviewer(Alice)
Teacher(Alice)
 $\exists y_1 \text{ reviews}(Alice, y_1)$

Query Rewriting

$\exists x \exists y \text{ Teacher}(x) \wedge \text{reviews}(x, y)$
 $\exists x \text{ Professor}(x) \wedge \text{Reviewer}(x)$
 $\exists x \text{ Teacher}(x) \wedge \text{Reviewer}(x)$
 $\exists x \exists y \text{ Professor}(x) \wedge \text{reviews}(x, y)$

A Use of Ontologies for Grouping Answers

http://dbpedia.org/resource/Mitropa_Cup	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Canadian_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Women's_National_Basketball_Association	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/AF2	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Arena_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Scottish_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Atlantic_League_of_Professional_Baseball	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/National_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/American_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Western_Baseball_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/World_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/United_States_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Canadian_Baseball_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Frontier_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Northern_League_(baseball,_1993%E2%80%A932010)	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Pacific_Coast_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Eastern_League_(baseball)	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Southern_League_(baseball)	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Texas_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Carolina_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Midwest_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Northwest_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/American_Southwest_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague

Figure: Extracts of the 3479 results of Sports League from DBpedia.

Structuring the Answer Set

Our aim is to present the **complete** set of answers to the user in an **interesting** and **semantically graspable** way.

A Use of Ontologies for Grouping Answers

http://dbpedia.org/resource/Mitropa_Cup	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Canadian_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Women's_National_Basketball_Association	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/AF2	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Arena_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Scottish_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Atlantic_League_of_Professional_Baseball	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/National_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/American_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Western_Baseball_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/World_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/United_States_Football_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Canadian_Baseball_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Frontier_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Northern_League_(baseball,_1993%E2%80%932010)	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Pacific_Coast_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Eastern_League_(baseball)	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Southern_League_(baseball)	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Texas_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Carolina_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Midwest_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Northwest_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague
http://dbpedia.org/resource/Amateur_League	http://www.v3.org/1999/02/22-rdf-syntax-ns#type	http://dbpedia.org/ontology/SportsLeague

Figure: Extracts of the 3479 results of Sports League from DBpedia.

The semantic for grouping answers and the visibility of the complete set of answers makes our approach **different from top k** and from **clustering**.

Rewriting Directed Acyclic Graph

Structuring the Answer Set

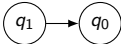
We define a directed acyclic graph of rewriting queries for semantically structuring the set of answers.

This graph is built following three steps :

- ▶ structure from rewriting queries
- ▶ saturation of entailments
- ▶ simplifications

Structure of Rewriting Queries

Building of an edge

- ▶ **query** : $q_0(x) := \exists y \text{ Teacher}(x) \wedge \text{reviews}(x, y)$
- ▶ **rule** : $\forall x \text{ Professor}(x) \rightarrow \text{Teacher}(x)$
- ▶ **rewriting query** : $q_1(x) := \exists y \text{ Professor}(x) \wedge \text{reviews}(x, y)$
- ▶ **resulting edge** : 

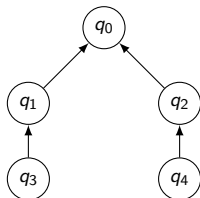
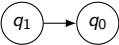


Figure: An example of rewriting graph.

Structure of Rewriting Queries

Properties

Entailment and Rewriting

 implies that q_1 entails q_0 .

Finite Set of Rewriting Queries

We use RDFS rules, so the set of rewriting queries is always finite.

Saturation of Entailments

We add edges to the rewriting graph for representing all entailments between queries.

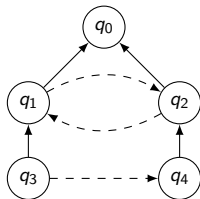


Figure: Rewriting graph completed with entailments (dashed edges)

Simplifications

Merging Equivalent Queries

We merge equivalent queries of the completed rewriting graph.

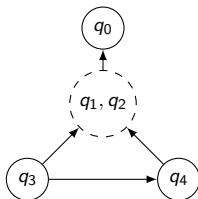


Figure: The equivalent queries q_1 and q_2 are merged.

The resulting graph is a directed acyclic graph.

Simplifications

Keeping the transitive reduction

We keep the transitive reduction of the graph.

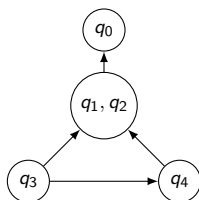


Figure: The edge from q_3 to q_1, q_2 is removed, because of the existing path through q_4 .

The transitive reduction is unique because the graph is acyclic.

Simplifications

Keeping the transitive reduction

We keep the transitive reduction of the graph.

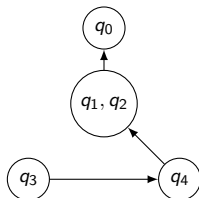


Figure: The edge from q_3 to q_1, q_2 is removed, because of the existing path through q_4 .

The transitive reduction is unique because the graph is acyclic.

A Grouping Method : The Cardinal-based Labellisation

Expectations on the answer groups

We aim to find semantically defined groups of answers of a query q , such as :

- ▶ each group's cardinality is close to a fixed integer t .
- ▶ the union of group's answers is the answers of q .
- ▶ the semantic definition of group is not entailed by a other one.

Coloring of the Rewriting Graph

We color the queries in the DAG of rewriting queries :

- ▶ in **black** if the rewriting query has more than t answers.
- ▶ in **red** otherwise.

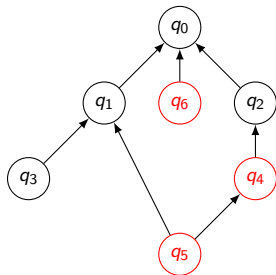


Figure: A example of coloring of a DAG of rewriting queries.

Grouping Answers from Colored Graph

Border of a Colored Graph

We select a set of rewriting queries and takes their answers as a semantically defined group of answers. The groups satisfy the expected properties.

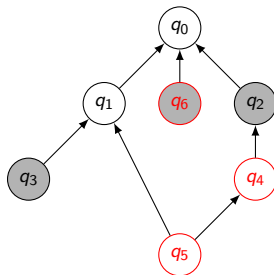


Figure: Black queries without black rewriting query are chosen and red queries which are not a rewriting of a red query or a chosen black query are chosen.

Conclusion

We proposed an approach to tackle the problem of apprehending the answers of an ontology-based query by grouping it semantically.

Our approach follows the following steps :

- ▶ building the graph of rewriting queries of the original query,
- ▶ enriching and simplifying this graph,
- ▶ extract of this graph queries such as their answers form semantic groups of answers.