

Short Time Scientific Mission Report

COST-STSM-IC1302-36991

Andrea Calí and Martín Ugarte

April 20, 2017

Purpose of the STSM. The main objective of this STSM was to study possible applications of recent techniques developed in the area of Dynamic Query Processing [2] to the field of Keyword Search, specifically in the context of Deep Web [1]. Intuitively, when posting a keyword query to Deep Web sources one needs to process a recursive query plan in order to comply with the access limitations of such sources. While previous research has addressed the problem of pruning such recursive plans, we focused on indexing and caching techniques for processing recursion as updates to a database.

Description of the work carried out during the STSM. We addressed to main lines of research:

1. The application of the Dynamic Yannakakis Algorithm [2] to the context of Deep Web query processing.
2. In the context of caching related to Deep Web query processing, we addressed the problem of query containment from the foundational point of view.

Description of the main results obtained. We obtained the following results for the two lines of research addressed:

1. We proposed a novel framework for efficiently processing recursive query plans under access limitations by using dynamic caching of partial query results satisfying join dependencies.
2. We established tight complexity bounds for the problem of query containment in the context of Deep Web. We proved that in the case of binary predicates the problem is Π_2^p -complete; our lower complexity bound shows that the problem is already Π_2^p -hard for the case of a single domain with input-only predicates. As a corollary, we proved that conjunctive query containment (in the classic sense) with the domain as part of the input is Π_2^p -complete.

Future collaboration with the host institution.

1. We plan to establish complexity bounds for showing that the techniques for dynamic processing of recursive query plans are optimal in the worst case. In particular, we would ideally achieve linear precomputation and constant-delay enumeration for free-connex acyclic conjunctive queries under access limitations.

2. In the same context of dynamic query processing, we plan to validate our results by implementing and testing a prototype.
3. For the theoretical line of research, we plan to address further relevant cases of query containment. For instance different query languages (e.g. conjunctive queries with inequalities) or integrity constraints on the data.
4. As an application of our theoretical results, we conjecture that the problem of deciding if a regular path query is contained in a conjunctive query is also Π_2^P -complete.

Foreseen publications/articles resulting from the STSM. We already submitted a short paper to the Alberto Mendelzon Workshop on Foundation of Data Management (AMW 2017), where we discuss some of our theoretical findings. We plan to submit a framework paper to the third Keystone conference (IKC 2017) presenting our techniques for incremental evaluation of recursive queries. Later on we will submit a paper to the VLDB journal about our theoretical findings and their applications. Finally, we plan to submit a PVLDB paper on dynamic query processing in the Deep Web.

References

- [1] A. Calì and D. Martinenghi. Querying data under access limitations. In *Proc. of ICDE 2008*, pages 50–59, 2008.
- [2] M. Idris, M. Ugarte, and S. Vansummeren. The dynamic yannakakis algorithm: Compact and efficient query processing under updates. In *Proc. of SIGMOD 2017*, 2017. To appear.



Andrea Calì
Senior Lecturer
Department of Computer Science and Information Systems
Birkbeck, University of London
Malet Street
London WC1E 7HX
United Kingdom
andrea@dcs.bbk.ac.uk

20th April, 2017

CONFIRMATION OF RESEARCH VISIT — REPORT OF HOST INSTITUTION FOR STSM OF COST ACTION IC1302, DR MARTÍN UGARTE; STSM REF. COST-STSM-IC1302-36991.

This is to certify that Dr Martín Ugarte, of the Université Libre de Bruxelles, recently completed a research visit at the Department of Computer Science and Information Systems, University of London, Birkbeck College (UK), from the 10th to the 20th of April, 2017.

I further confirm that during the visit Dr Ugarte and I have started a research collaboration on efficient query processing in the Deep Web.

I confirm that this STSM was very successful and that it opened a fruitful collaboration between the participating countries within the framework of COST Action IC1302 *Keystone*. With Dr Ugarte we already submitted a short paper; we do plan to continue the line of research that we just started and we are confident that this will yield further high-quality research results to be published in top venues.

Yours faithfully,

Andrea Calì

A handwritten signature in blue ink, appearing to read "Andrea Calì", written in a cursive style.