

Internship Proposal

Enumerating Query Results on Multitrees

Antoine Amarilli <antoine.amarilli@telecom-paris.fr>

Louis Jachiet <louis.jachiet@telecom-paris.fr>

Luc Segoufin <luc.segoufin@inria.fr>

Background. When evaluating queries over large databases, the result is often too large to be fully materialized. In such cases, one can study different tasks, e.g., counting the number of solutions, testing efficiently whether an input candidate is in the set of solutions, or enumerating solutions one after the other. This internship focuses on the enumeration problem, formalized in the framework of enumeration algorithms¹.

Enumerating query results on databases has already been studied in many research works, and there are several known classes of query languages and databases for which enumeration can be performed with linear-time preprocessing and constant delay between answers. For instance, this can be achieved for queries in first-order logic on bounded-expansion structures [2] and for monadic second-order queries on structures of bounded treewidth [3, 1].

Research topic: The internship will focus on a class of structures called multitrees² which are a natural generalisation of trees. While trees only have a single root, multitrees can have multiple roots which can share descendants; but they still ensure that for any choice of a root and a node, there is at most one path going from the root to the node.

As explained above, on trees, it is possible to enumerate the results of monadic second-order queries with linear-time preprocessing and constant delay. This result no longer seems to hold for multitrees. The goal of the internship is then to determine which queries can be enumerated with linear-time preprocessing and constant-delay on multitrees.

Supervision and environment. The internship will take place from March to August, 2021, at the École normale supérieure, 45 rue d'Ulm, 75005 Paris (VALDA team). It will be supervised by Antoine Amarilli and Louis Jachiet (Télécom Paris) and Luc Segoufin (Inria Paris).

References

- [1] A. Amarilli, P. Bourhis, L. Jachiet, and S. Mengel. A Circuit-Based Approach to Efficient Enumeration. In *ICALP*, 2017.
- [2] W. Kazana and L. Segoufin. First-order query evaluation on structures of bounded degree. *LMCS*, 7(2), 2011.
- [3] W. Kazana and L. Segoufin. Enumeration of monadic second-order queries on trees. *TOCL*, 14(4), 2013.

¹https://en.wikipedia.org/wiki/Enumeration_algorithm

²<https://en.wikipedia.org/wiki/Multitree>