

Flexible Querying of Network-Structured Data

Project Aims

Current proposals for languages to query network-structured data both lack the power to express certain queries that are useful in envisaged applications and provide only limited capabilities for flexible querying, with no ability to rank the answers for users.

We aim to develop a theoretical framework within which the properties of languages for flexible querying of network-structured data can be investigated. We also plan to implement and evaluate a prototype tool specifically for the flexible querying of RDF.

Application areas and examples

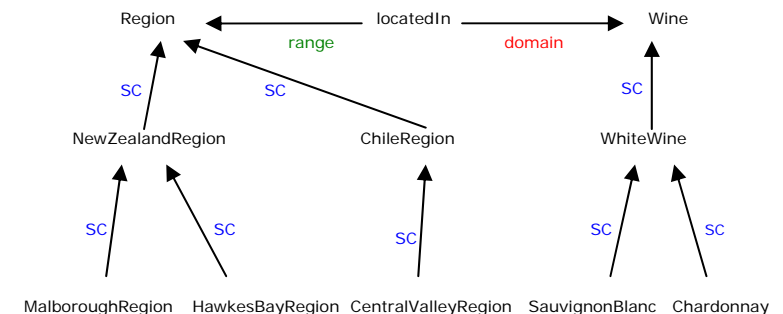
Examples of application areas in which information has a network structure include the World Wide Web, geographical information systems, and social networks. Users in these areas are often interested in the ways in which the data items relate to one another, in other words, in how they are connected. For example, Nokia Research are experimenting with finding connections in data relating to mobile phone users in order "to form spontaneous collaborations between mobile devices and their users".

Queries like the above can be difficult to pose with current solutions. Furthermore, if a query returns no answers or returns fewer answers than a user expects, it is up to the user to reformulate the query in order to find more answers.

Flexible querying provides a solution to the problems raised above. A flexible query system can reformulate queries automatically in response to either no answers being found for an original query or the user requesting further answers. Such a system could also return answers to the user ranked in an order that reflects how "closely" the answers matched

the original query, as well as informing the user which reformulated query they match.

For example, suppose that we are querying RDF data about wines, and we want to find cheap Sauvignon Blanc wines from the Marlborough Region in New Zealand. Some domain knowledge about wines is represented as a very simple RDFS ontology in the diagram below.



If there are no wines that satisfy our requirements, a flexible query processing system can relax our constraints using information from the ontology. Since Sauvignon Blanc is a type of white wine (indicated by the arc labelled SC) and the Marlborough Region is a New Zealand Region, the system can automatically reformulate our original query to ask for cheap white wines from the Marlborough Region or for cheap Sauvignon Blanc wines from the New Zealand Region.

Publications

- [1] C.A. Hurtado, A. Poulouvassilis and P.T. Wood, "A Relaxed Approach to RDF Querying," in *Proc. 5th Int. Semantic Web Conference*, LNCS 4273, Springer-Verlag, 2006, pp. 314-328.
- [2] C.A. Hurtado, A. Poulouvassilis and P.T. Wood, "Query Relaxation in RDF," *Journal on Data Semantics X*, LNCS 4900, Springer-Verlag, 2008, pp. 31-61.



Project Leaders

Peter Wood

ptw@dcs.bbk.ac.uk

Carlos Hurtado (Univ. Chile)

churtado@dcc.uchile.cl

Other Project Members

Alex Poulouvassilis

ap@dcs.bbk.ac.uk

Claudio Gutierrez (Univ. Chile)

cgutierr@dcc.uchile.cl

Project Details

Funded by the Royal Society

May 2007 – April 2009

Project Web Site

www.dcs.bbk.ac.uk/~ptw/query.html