Explaining Knowledge Bases for Query Answering on the Semantic Web

Deborah L. McGuinness
Knowledge Systems Laboratory
Stanford University, Stanford, CA
dlm@ksl.stanford.edu

Through the rapid growth of the World Wide Web, there now exist seemingly limitless stores of information. These information stores are increasingly at least semi-structured in nature, and many times are actually structured knowledge bases of objects. Search engine use has also seen increased usage and has become the source of numerous articles both about search's growing importance as well as about the increasing user demands for enhanced techniques for object manipulation and access(e.g., [Hagen, et al, 2000]). Many of today's applications require "real" query answering instead of simply depending upon search and document retrieval. Application agents and users still want access to the document(s) from which an answer is obtained but they also need the option of receiving an explanation of why some object has been returned as answer and need object manipulation options. This is required for interoperability and for trust. In this work, we will explore the notion of an explanation for such query answering on the emerging semantic web and posit a foundational strategy for explanation.

We view a baseline query answering notion to include the ability to return identified information that satisfies a query specification. Document retrieval does not satisfy this definition as query answering since it does not identify the portion of the document (i.e. the document object(s)) that contains the answer. Today's agent or human user who asks for universities in Santa Clara County for example, does not just want documents including "Stanford University" or "UCSC" (or the university name and its address), they want the identified object representing Stanford University or UCSC, along with the option of accessing object properties such as the address, the county of the object, the source of that information, relationships between the object and other objects such as the university's student population, etc.

Today's query answering does not stop with just the above baseline notion however. We believe today's query answering requires an optional justification that gives an agent or user the option of receiving computationally operational information about why an object is believed to meet a query specification. We refer to this as the answer justification and state that a minimal justification includes the inference rule application(s) along with bindings of all variables instantiated in the inference rule(s). We choose this view of justification building on the notion of explanation for description logics (e.g., [McGuinness, 1996; Borgida, et. al, 1999]).

A simple inference rule is "told" or stated information. A justification for Stanford University being a University in Santa Clara County may be that the information was stated on the Stanford University Home Page and it was stated information in knowledge base XYZ. A user or agent may want to view the source of information and they may need access to the sources of information that included the stated information. A more deductive justification may be that the query answerer used a geographic containment inference rule along with the bindings, ?city=Stanford and ?region=Santa Clara County. The geographic containment rule may state that ?x is in ?region if ?x is in ?city and ?city is in ?region. The user may also want to see source information that Stanford University is located in the city Stanford, as stated on the Stanford University home page and Stanford is a town in Santa Clara County as stated on the Santa Clara county government home page. A user may wish to "believe" this answer because she believes the sources and the inference rule.

In our work, we start with the notion of justifications for deductions in knowledge representation and reasoning systems such as the DAML+OIL description logic-based system and explore the notions of explanation for query answering.

References

[Borgida,et.al, 1999] Borgida, Franconi, Horrocks, McGuinness, and Patel-Schneider. "Explaining ALC subsumption". Proceedings of the International Workshop on Description Logics, pp 33-36, Linköping, Sweden, July 1999.

[Hagen, et.al, 2000] Hagen, Manning, and Paul. "Must Search Stink" Forrester Research TechStrategy Briefing, June 2000.

[McGuinness, 1996] McGuinness. "Explaining Reasoning in Description Logics". Ph.D. Thesis, Rutgers University, 1996. Technical Report LCSR-TR-277.