Chapter XV

A Semantic Web Based Approach for Context–Aware User Query Formulation and Information Retrieval

Hanh Huu Hoang  
Vienna University of Technology, Austria

Tho Manh Nguyen  
Vienna University of Technology, Austria

A Min Tjoa  
Vienna University of Technology, Austria

ABSTRACT

Formulating unambiguous queries in the Semantic Web applications is a challenging task for users. This article presents a new approach in guiding users to formulate clear requests based on their common nature of querying for information. The approach known as the front-end approach gives users an overview about the system data through a virtual data component which stores the extracted metadata of the data storage sources in the form of an ontology. This approach reduces the ambiguities in users’ requests at a very early stage and allows the query process to effectively perform in fulfilling users’ demands in a context-aware manner. Furthermore, the approach provides a powerful query engine, called context-aware querying, that recommends the appropriate query patterns according to the user’s querying context.
MOTIVATION

The Semantic Web and ontologies have created a promising background for applying the intelligent techniques in information systems, especially in personal information management (PIM) systems. In PIM systems, effectively retrieving information from a huge amount of data of an individual is a challenging issue. The virtual query system (VQS) (Hoang, Andjomshoaa, & Tjoa, 2006) of the SemanticLIFE framework is an approach of using semantic Web techniques with a user-oriented method in order to tackle this challenge.

The SemanticLIFE project (Ahmed, Hoang, Karim, Khusro, Lanzenberger, Latif et al., 2004) is an effort to realize Vanevar Bush’s vision of Memex (Bush, 1945) by providing a general semantic PIM system. The SemanticLIFE system integrates a wide variety of data sources and stores them in an ontological repository. In the VQS-enhanced SemanticLIFE, the user is supported in issuing imprecise queries to retrieve the rich semantic information from the user’s historical personal data. However, users themselves often do not actually know or remember the specific qualities of what they are looking for, but have some awareness of other things related to the desired items (Quan, Huynh, & Karger, 2003). The VQS supports users in this nature when querying the information from the huge ontological repository effectively not only in the initial phase with offered “virtual information” but during the query process with the “context-based” querying features.

Furthermore, as mentioned above, the user’s nature of asking questions is that the user often does not know what the user is looking for; but the user remembers some concepts about related information the user is looking for. This leads us to a way of querying (browsing/navigating) the system by using redefined query templates (patterns) based on the user’s querying context. This would help the user not to be embarrassed in a new phase of query formulation.

The difficulty of query formulation appears not only in the initial phase but it continues in the query process or query refinement. During the query process, the user is asked for new requests using the new knowledge to get the information of interest. In order to ease the user from thinking of new constraints of their queries, we propose a new way based on the users’ nature, that is, preferring to customize the query patterns to make new queries. We trace the context of the user’s query process and recommend to the user the appropriate query patterns matching up the user’s query context.

Our approach originates from the user-side manner in trying to formulate unambiguous requests as early as possible during the querying process. The principle of the approach follows “better known, clearer request” and “customizing than creating.” If users are aware of what information they possess, they could ask precise queries against their stored data. This helps the query refinement process of the system by eliminating ambiguities at a very early stage of the query process. These approaches resulted in our query system, the VQS for the SemanticLIFE framework, with a query language called the virtual query language, and with a new way of query formulation entitled pattern-based and context-aware querying process.

The remainder of this article is organized as follows. The related work to our research is mentioned in Section 2. An overview of the SemanticLIFE project and the VQS is underlined in Section 3. Section 5 describes the details of the “virtual data component.” Section 6 presents the VQS’s innovative feature for query formulation and information retrieval. Main points of the VQS implementation are pointed out in Section 8. A summarized example is presented in the Section 9. Finally, the article is concluded with a sketch of the intended future work in Section 10.