Guest editors' introduction to the special issue on flexible queries in information systems

Patrick Bosc · Allel Hadjali · Gabriella Pasi

Nowadays, with the increasing amount of technologies which support the storage of data in digital format, the problem of accessing the information relevant to specific needs becomes a ubiquitous challenge. In particular, with the widespread expansion of the Internet and of the World Wide Web, myriads of databases are available online, providing massive repositories which may be accessed to reply to ordinary users' queries. On the other hand, due to the lack of knowledge about the contents of the large Web repositories and to the limited capability of the rigid standard querying interfaces, often users may not know how precisely express their needs and they may formulate queries that lead to unsatisfactory results. To improve the user friendliness of querying interfaces, as well as, to enhance querying facilities, several researchers have addressed the need for advanced systems to support user access to available information.

One desired advanced aspect of such systems is that they should show some forms of flexibility by allowing query languages with soft or flexible constructs. Flexible queries are requests in which user's preferences can be easily expressed. The Fuzzy Set formal framework provides an appropriate tool for supporting the expression of users' preferences; the user may not only specify crisp conditions, but also *fuzzy* ones whose satisfaction becomes a matter of degree rather than true or false.

P. Bosc · A. Hadjali (✉)
IRISA/ENSSAT, Lannion Cedex, France
e-mail: hadjali@enssat.fr

P. Bosc
e-mail: bosc@enssat.fr

G. Pasi
DISCO, Milano, Italy
e-mail: pasi@disco.unimib.fr

Preferences may be expressed at two distinct levels: inside elementary conditions as well as in the aggregation of conditions. In the first case, the objective is to express that some values are more adequate than others (from total satisfaction to complete rejection), while in the second case, preferences are intended as levels of importance associated with the single conditions. This approach generalizes Boolean queries where a condition is satisfied or not, and all the conditions are equally important. The major interest for flexible queries lies in the fact that their answers are no longer a classical unordered set of elements, but they are indeed discriminated, thus reflecting the compliance with the preferences stated in the query. So it becomes possible to get the best k answers or those having a matching score over a given threshold.

This special issue focuses on recent developments about flexible querying in Databases Management Systems and Information Retrieval Systems. It was inspired by "the Seventh International Conference on Flexible Query Answering Systems (FQAS 2006)" held on June 2006 in Milan, Italy. This special issue consists of four contributions that discuss innovative ideas in the area of flexible querying.

The first paper, "Relaxing RDF Queries Based on User and Domain Preferences" by Dolog, Stuckenschmidt, Wache and Diederich, discusses how queries over heterogeneous RDF data can be rewritten using knowledge about the domain model and user preferences. The proposed approach combines query refinement and query relaxation in order to provide robust, personalized access to such data. A prototypical implementation of the system is then presented in the e-learning domain by using real world data about e-learning resources in computer science.

In the second paper, "Incremental Controlled Relaxation of Failing Flexible Queries" by Bosc, Hadjali and Pivert, the authors address the problem of handling failing queries in the context of flexible querying. A cooperative approach supported by a relaxation mechanism is described. Query relaxation relies on a parameterized tolerance relation that allows transforming a gradual predicate (represented by means of a fuzzy set) into an enlarged one. Controlling the relaxation process is directly stemmed from the fuzzy set-based semantics underlying the tolerance relation of interest.

The third paper, "A Fuzzy Extension for the XPath Query Language" by Campi, Damiani, Guinea, Marrara, Pasi and Spoletini, describes an approach to bring flexibility into query languages for XML. XML has become a widespread format for data exchange over the Internet and it is crucial that the system for querying XML data goes beyond Boolean selection, and allows the user to define vague and soft selection conditions in order to obtain ranked results according to her/his preferences. To do so, a framework, called FuzzyXPath, is proposed. Flexibility is supported in FuzzyXPath at two levels by allowing: fuzzy constraints in query formulation and partial matching between two XML trees.

Finally, the fourth paper, "Issues in the Practical Use of the OWA Operators in Fuzzy Querying" by Zadrozny and Kacprzyk, is another well-founded example of introducing flexibility in querying traditional databases. The authors propose an extended version of SQL supporting linguistic terms in queries. In particular, linguistic quantifiers such as "most" and "almost all" are considered. They play the role of flexible aggregation operators and are interpreted using Yager's OWA operator. The implementation of this study is illustrated on the authors' prototype FQUERY.

We would like to express our gratitude to the referees for their work during the review process as well as to the authors of all submitted papers. Special thanks are also due to Prof. Ras, co-editor-in-chief of this journal, for providing us the opportunity for editing this special issue.

Patrick Bosc, Allel Hadjali and Gabriella Pasi
Guest Editors