Chapter IX
Qualifying Objects in Classical Relational Database Querying

Cornelia Tudorie
University Dunarea de Jos, Galati, Romania

ABSTRACT

The topic presented in this chapter refers to qualifying objects in some kinds of vague queries sent to relational databases. We want to compute a fulfillment degree in order to measure the quality of objects when we search them in databases. After a discussion on various kinds of object linguistic qualification, with different kinds of fuzzy conditions in a fuzzy query, a new particular situation is proposed to be included in this subject: the relative object qualification as a query selection criterion, that is, queries with two conditions in which the first one depends on the results of the second one. It is another way to express the user’s preferences in a flexible query. In connection with this, a new fuzzy aggregation operator, AMONG, is defined. We also propose an algorithm to evaluate this kind of queries and some definitions to make it applicable and efficient (dynamic modeling of the linguistic values and unified model of the context). We demonstrate these ideas with software already implemented in our lab.

INTRODUCTION

Database querying by various selection criteria can often confront a major limitation: the difficulty to realize and express precise criteria for locating the information. This happens because people do not always think and speak in precise terms, or they do not have details on the data range. The research community recently proposed a new way to query databases, more expressive and flexible than the classical one. It is about vague queries, for example: “retrieve the persons well paid which live not too far from the office”, of course, formulated in an adequate query language. The main reason to use the vague predicates well paid and not too far is to express more flexibly the user’s preferences and at the same time to rank the selected tuples by a degree of criteria satisfaction.

When a precise criterion, like “salary > 500 and distance home-office < 200” is required, it may return an empty list, even if there are a lot of persons having attribute values very close to the specified ones. As well, the same precise criterion may return a complete list of all persons, without any helpful ordering. So, it would be useful to provide intelligent interfaces to databases, able to interpret and evaluate imprecise criteria in queries.
Some important advantages resulting from including vague criteria in a database query may be:

- Easy to express queries
- The possibility to classify database objects by selecting them based on a linguistic qualification
- The possibility to refine the result by assigning to each tuple the corresponding fulfillment degree (the degree of criteria satisfaction); in other words, to provide a ranked answer according to the user’s preferences.

Under this circumstance, when vague queries are accepted, fuzzy set membership functions are convenient tools for modeling the user’s preferences in many aspects.

Fuzzy sets theory (Bouchon-Meunier, 1995; Dubois, Ostasiewicz, & Prade, 1999; Yager & Zadeh, 1965) is accepted as one of the most adequate formal frameworks to model and to manage vague expressions. Two research areas are important for the fuzzy theory applied in database field: fuzzy querying regular databases and storing fuzzy information in databases. There are many scientific works regarding database fuzzy querying: general reference books (e.g., Galindo, Urrutia, & Piattini, 2006), but also many articles in journals and communications at conferences. Some of them propose fuzzy extensions of the standard query language for relational database (SQL), able to interpret and evaluate fuzzy selection criteria; others proposed intelligent interfaces to fuzzy querying classical databases. The most important of those included:

- SQLf (Bosc & Pivert, 1995; Goncalves & Tineo, 2001a, 2001b; Projet BADINS, 1995, 1997) and FSQL (Galindo et al., 2006) are extensions of the SQL language, allowing flexible querying.
- FQUERY (Kacprzyk & Zadrozny, 1995, 2001) and FuzzyBase (Gazzotti, Piancastelli, Sartori, & Beneventano, 1995) are fuzzy querying engine for relational databases.

In this book, the reader can find a chapter by Urrutia, Tineo, and González, studying SQLf and FSQL languages. There is also another chapter including a review about flexible querying and it has been written by Kacprzyk, Zadrożny, de Tré, and de Caluwe. Other works have developed new data models able to taking into account imperfect information. Fundamental contributions have been made by Buckles and Petry (1982); Medina, Pons, and Vila (1994); and Prade and Testemale (1984). Galindo et al. (2006) also define a running fuzzy database, which stores imperfect data represented by fuzzy possibilistic distributions, fuzzy degrees, and so forth.

From the beginning, it is important to remark that we deal in this chapter with relational database fuzzy querying (fuzzy queries on crisp data), but not fuzzy database querying (queries on fuzzy data). The first presented items are already talked about in Project BADINS (1995, 1997); Bosc and Pivert (1992); Bosc and Prade (1997); Dubois and Prade (1996); Kacprzyk and Zadrozy (2001); and many others, but we consider that it is useful to rediscuss them in order to propose an original classification of the various kinds of object linguistic qualifications. In this context, the relative object qualification will be proposed as a new kind of selection criteria, and it will be included in this classification.

Some practical examples inspired us to develop this study. Let us compare the queries:

*Retrieve the cars having the speed greater than 240*

*Retrieve the cars having high speed*

*Retrieve the inexpensive and high speed cars*

*Retrieve the inexpensive cars among the high speed ones*

They are increasingly more complex: they start with a classical crisp query and go to more complex queries, including vague terms in the selection.