Chapter VI

On the Versatility of Fuzzy Sets for Modeling Flexible Queries

P. Bosc
IRISA / ENSSAT – Université de Rennes 1, France

A. Hadjali
IRISA / ENSSAT – Université de Rennes 1, France

O. Pivert
IRISA / ENSSAT – Université de Rennes 1, France

ABSTRACT

The idea of extending the usual Boolean queries with preferences has become a hot topic in the database community. One of the advantages of this approach is to deliver discriminated answers rather than flat sets of elements. Fuzzy sets are a natural means to represent preferences, and many works have been undertaken to define queries where fuzzy predicates can be introduced inside user queries. The objective of this chapter is to illustrate the expressiveness of fuzzy sets with the division operator in the context of regular databases. Like other operators, the regular division is not flexible at all and small variations in the data may lead to totally different results. To counter this behavior, a variety of extended division operators founded on fuzzy sets are suggested. All of them obey a double requirement: to have a clear meaning from a user point of view and to deliver a resulting relation which is a quotient.

INTRODUCTION

The idea of introducing preferences into queries is gaining more and more attention in the database community. After some initial works in the 1970s and 1980s, such as nearest neighbors (Friedman, Baskett, & Shustek, 1975), Deduce2 (Chang, 1982), Preferences (Lacroix & Lavency, 1987), Ares (Ichikawa & Hirakawa, 1986), and Vague (Motro, 1988), a new stream of work, including top-k queries (Bruno, Chaudhuri, & Gravano, 2002), PreferenceSQL (Kießling & Köstler, 2002), Skyline (Börzsönyi, Kossmann, & Stocker, 2001), and the synthesis in Chomicki (2003), is based on the use of preferences inside user queries. Undoubtedly, this contributes to make queries (and database systems) more flexible, but it must be noticed that most (not to say all) of these approaches or systems do not call on fuzzy sets. We believe that the main advantage for founding flexible queries on fuzzy sets lies in the generality of the approach, which allows notably the combination of preferences.
over different attributes (or items). Whatever the approach taken, the final objective is no longer to specify acceptable elements, but rather to define an order (partial or total) over the elements retrieved from the database.

This chapter is situated in the context of regular relational databases, and its main objective is to illustrate the expressive power of fuzzy sets for modeling and interpreting a wide range of division operations involving preferences. Several works dealing with the expression and interpretation of fuzzy queries have been carried out (in particular, Tahani, 1977; Kacprzyk & Ziolkowski, 1986; Bosc & Pivert, 1995). Selection, projection, Cartesian product, join, as well as set-oriented operations have been defined so as to take preferences into account. It turns out that these operations can be extended only in a straightforward manner. On the contrary, the division operation is a bit more complex, and it gives birth to a bunch of extensions in the context of fuzzy sets. The relational division has been the subject of some research works in the fuzzy set community in the past with two types of objectives: (1) to extend the operation in the presence of fuzzy relations (Bosc, Dubois, Pivert, & Prade, 1997; Cubero, Medina, Pons, & Vila, 1994; Dubois & Prade, 1996; Mouaddib, 1993; Yager, 1991) and (2) to define the division when the attributes of the relations may be imprecise (Dubois, Nakata, & Prade, 2000; Galindo, Medina, Cubero, & Garcia, 2001; Umano & Fukami, 1994). This chapter falls in the first category where data are precisely known and different types of division operators are studied in the context of flexible queries addressed to regular databases: division of fuzzy relations and approximate division intended for some tolerance to exceptions.

Let us recall that the relational division is somewhat analogous to the integer division. In effect, similarly to the integer division which returns an integer subject to a constraint (the largest one such that its product with the divisor is smaller than or equal to the dividend), the relational division returns a relation subject to an inclusion constraint. As a consequence, one key point in extending the division is the assessment of the quality of quotient of the result delivered. In the following, we aim at illustrating how the division can become a flexible operator with two central concerns: (1) to suggest extensions which have a clear semantics from a user point of view and (2) to show that the extended operator still has the characteristic of a division, that is, that it delivers a quotient.

The rest of the chapter is structured as follows. The second section is devoted to an overview of flexible queries in the context of fuzzy sets. Both an algebraic framework and the key elements of an SQL-like query language are outlined. In the third section, basics on the relational division are recalled and some motivations for extending this operator to make it more flexible are pushed forward. The fourth section deals with the division of fuzzy relations, that is, relations whose tuples are graded to express the extent to which each of them is compatible with the fuzzy concept represented by the considered fuzzy relation. Different lines of extension are proposed and the fact that a quotient is delivered is the criterion for their acceptability. In the next section, the notion of exception in the context of a division is introduced. Quantitative and qualitative exceptions are distinguished, which are the source of two families of approximate divisions. In the sixth section, an approximate division based on the idea of erosion (reduction) of the divisor and the dilation (augmentation) of the dividend is suggested. More precisely, erosion and dilation are performed on the basis of a parameterized proximity relation. Finally, the conclusion summarizes the contributions of the chapter and draws some lines of future work.