Chapter 1
Fuzzy Database Modeling: An Overview and New Definitions

Angélica Urrutia
Universidad Católica del Maule, Chile

José Galindo
Universidad de Málaga, Spain

ABSTRACT

In this chapter the authors present an overview of different fuzzy database modeling definitions by different authors. They highlight the FuzzyEER model, which is an extension of an EER model incorporating fuzzy semantics and notation. FuzzyEER is probably the most complete modeling tool. It has numerous characteristics, and in this chapter the authors list the main components and enhance some definitions relative to fuzzy attributes, fuzzy degrees associated to one or more attributes or with an independent meaning, fuzzy entities, as well as the definition of multivalued disjunctive attributes.

INTRODUCTION

On occasions the term “imprecision” embraces several meanings between which we should differentiate. For example, the information we have may be incomplete or “fuzzy” (diffuse, vague), or we may not know if it is certain or not (uncertainty), or perhaps we are totally ignorant of the information (unknown), we may know that that information cannot be applied to a specific entity (undefined), or we may not even know if the data can be applied or not to the entity in question (“total ignorance” or value “null”) (Umano and Fukami, 1994). Each of these terms will depend on the context in which they are applied and these concepts have been widely studied in the database context in many papers, such as (Galindo et al., 2006; Galindo, 2008; Zadrożny et al., 2008).

The management of uncertainty in database systems is a very important problem (Motro, 1995) as the information is often vague. Motro states that fuzzy information is content-dependent, and he classifies it as follows:

DOI: 10.4018/978-1-60566-814-7.ch001
Fuzzy Database Modeling

• **Uncertainty**: We cannot know whether the information is true or false. For example, “John may be 38 years old”.

• **Imprecision**: The information available is not specific enough. For example, “John may be between 37 and 43 years old”, — disjunction — “John is 34 or 43 years old”, — negative — “John is not 37 years old”, or even unknown.

• **Vagueness**: The model includes elements (predicates or quantifiers) which are inherently vague, for example, “John is in his early years”, or “John is at the end of his youth”. However, once these concepts have been defined, this case would match the previous one (imprecision).

• **Inconsistency**: It contains two or more pieces of information, which cannot be true at the same time. For example, “John is 37 and 43 years old, or he is 35 years old”; this is a special case of disjunction.

• **Ambiguity**: Some elements of the model lack a complete semantics (or a complete meaning). For example, “It is not clear whether they are annual or monthly salaries”.

**Zadeh** (1965) introduces the fuzzy logic in order to deal with this type of data. Traditional logic, because it is bi-valued, can only operate with concepts like: yes or no, black or white, true or false, 0 or 1, which allowed just for a very limited knowledge representation. Although there are other logics which take more truth values, namely multi-valued logics, fuzzy logic is one extension which takes endless truth levels (or degrees), associating the concept of membership degree or truth degree in an interval $[0,1]$ within the fuzzy logic theory.

**Fuzzy databases** have also been widely studied (Galindo, 2008), with little attention being paid to the problem of conceptual modeling (Chaudhry et al., 1999) and focusing the research mainly in fuzzy queries (Zadrożny et al., 2008). This does not mean that there are no publications, however, but that they are sparse and with no standard. Therefore, there have also been advances in modeling uncertainty in database systems (Buckles and Petry, 1985; Kerre and Chen, 1995; Chen, 1998; Yazici and George, 1999) including object-oriented database models (Van Gyseghem et al., 1993; George et al., 1996; Caluwe, 1997; Bordogna et al., 1999; Yazici and George, 1999; Ma et al., 2004). Probably, the most complete approach was published in (Galindo et al., 2006) in the so-called FuzzyEER model.

At the same time, the extension of the ER model for the treatment of fuzzy data (with vagueness) has been studied in various publications (Zvieli and Chen, 1986; Ruspini, 1986; Vandenberghe, 1991; Chaudhry et al., 1994 and 1999; Chen and Kerre, 1998; Chen 1998; Kerre and Chen, 2000; Vert et al., 2000; Ma et al., 2001), but none of these refer to the possibility of expressing constraints by using the tools by fuzzy sets theory. In (Kerre and Chen, 1995) a summary of some of these models can be found. On the other hand, the main methodologies of databases design, such as (Elmasri and Navathe, 2000), have not paid attention to the modeling of data with uncertainty, although the intent of uncertainty modeling of the real world is rarely absent.

Based on these concepts, in this chapter we will discuss different approaches, by various authors, related to the uncertainty conceptual modeling problem in database models. After, we summarize the FuzzyEER model, a tool for fuzzy database modeling with many advantages with respect to the previous modeling tools. Some of its characteristics are: fuzzy values in the attributes, degree in each value of an attribute, degree in a group of values of diverse attributes, as well as, fuzzy entities, fuzzy relationships, fuzzy aggregation, fuzzy constraints. Furthermore, we include here some new definitions for the FuzzyEER model.