4. LOGICAL DATABASE MODELS FOR UNCERTAIN DATA

Abstract: In recent years, a primary objective of the database community has been the incorporation of structured and complex data types. This has led to new database models based on both the established relational paradigm as well as the object-oriented paradigm. These new database models are the non-1NF relational data model (also called NF² data model), object-oriented data model, and deductive object-oriented data model. These models are considered appropriate for modeling many non-traditional applications, such as CAD/CAM, imagery, multimedia, meteorology, geographic information systems, oceanography, etc. The NF² data model is better suited for office automation systems. On the other hand, object-oriented databases are more appropriate for CAD/CAM, multimedia database and geographical, meteorological, and oceanographic applications. Object-oriented databases coupled with logic and, as the more integrated approach, deductive object-oriented databases are more appropriate for knowledge intensive applications such as expert database systems. In this chapter, we introduce the logical database models such as the extended NF² data model, the fuzzy object-oriented data model, and the fuzzy deductive object-oriented data model to deal with complex information and uncertainty that arise in non-conventional applications.

Keywords: Complex Values, Fuzziness and Uncertainty, Object-Oriented Databases, Deduction, Querying Databases

4.1 Introduction

Despite the progress of the last decade in more fully utilizing structured data and, by such means, partially closing the paradigm mismatch between databases and computer languages, database technology has yet to make significant inroads into the targeted fields. This can, in part, be explained in terms of the lag time that is inherent to any new technology. It may also be that there is yet something lacking in the new models that limits their applicability. One thing is a coherent strategy for managing uncertainty since most effort in this direction has been performed under the assumptions of the older database paradigms [6,23,47,57,68,83] or knowledge-based systems [62,79]. We present a strategy here that is founded on several concepts relevant to uncertainty, i.e., fuzzy set theory.
Most of the existing database models are designed under the assumption of precise data storage and retrieval. Everything uncertain must be precisely defined; therefore, anything uncertain must be either precluded or made artificially precise. In real life, this assumption of precise world is often not appropriate for many complex applications (i.e. office automation systems, decision support systems, geographic and environmental information systems, expert database, multimedia database systems) since they involve in not only complex data/knowledge but also some kinds of uncertainty. When precise information is unavailable about the miniworld, it is often the case that some relevant information is available. In these cases, it may be advantageous to design methods for impreciseness by which this information can be stored, manipulated and retrieved. Uncertainty may also be present in requests to retrieve data, when users formulate their queries with imprecise terms.

In general, information is both complex, and imprecise or fuzzy when representing personalities, physical features of individuals, subjective opinions and judgments concerning medical diagnosis, economic forecasting, or personal evaluation, and in many other knowledge-intensive applications. Also, in natural languages, numerous quantifiers (e.g., many, few, some, almost, etc.) are used when conveying vague information. Apart from complexity, there have also been advances in modeling uncertainty in database systems [23,57,72,74,75], but most have been in context of the relational model [15,57,81]. Many of the existing approaches dealing with uncertainty are based on the theory of fuzzy sets [15,57].

The issues are handling different types of uncertainty and are dealing with it at both the conceptual and logical levels. The different types of uncertainty of complex objects at logical level considered here are as follows: First, consider the kind of uncertainty for which the true value is known to be one of a specific set of values. For example, a set of values of a Salary attribute might be 20,000-25,000, and the interpretation is purely disjunctive. Consider another kind of uncertainty for which the true value is not known (unk), does not exist (dne), or no information (ni) on whether a value exists or not. All can be considered a kind of null value. Another kind of uncertainty is that the information is available, but in descriptive and imprecise form. Such uncertain data is referred to as fuzzy. For example a value of an Age attribute is young.

In this chapter, uncertainty in non-conventional applications that is arising from many sources is represented via each one of the logical database models such as an extended NF² data model [57,73], object-oriented data models [23,74,84], and deductive object-oriented data models [40,41,42,45]. The chapter is organized as follows: Section 4.1 discusses the NF² database model. In Section 4.2, we describe the similarity-based fuzzy object-oriented database model. Section 4.3 presents how the fuzzy object-oriented data modeling is coupled with