Preface

In 1965 at the University of California, Berkeley, also called the "Athens of the Pacific," Lotfi A. Zadeh[1] introduced the theory of fuzzy sets and fuzzy logic, two concepts that laid the foundation of possibility theory in 1977. These terms were coined by him to deal with the phenomenon of vagueness, in the cognition process of the human being. According to Zadeh, "the theory of fuzzy sets is a step toward a rapprochement between the precision of classical mathematics and the pervasive imprecision of the real world… a rapprochement born of the incessant human quest for a better understanding of mental processes and cognition[2]."

Since then, an enormous quantity of congresses and publications around the world has intended to explore and develop this basic idea of vagueness and its industrial application. Zadeh also said: "at present, we are unable to design machines that can compete with humans in the performance of such tasks as recognition of speech, translation of languages, comprehension of meaning, abstraction and generalization, decision-making under uncertainty and, above all, summarization of information."

When we look at the growth of the Japanese industry in the 1980s we can understand the relevant impact of "fuzzy technologies" in the modeling and design of new products[3].

In these aspects, the gap between the industrial domain and the research domain can be seen in books, journals, articles, cases studies, proceedings, and so forth. In fact, these are the greatest tools to put the theoretical knowledge in action (i.e., in Idea Group Publishing you can find the latest advance in the research of information science, technology, and management). But it is diffi-

cult to find a pedagogical book to help the learning process of the students in computer science in the area of fuzzy databases.

I am glad to tell you that the book you have in yours hands has the courage to attack the problem of fuzzy databases, with a clear and direct approach guiding the reader step-by-step through the understanding process. Indeed this book has the ability to help you in the modeling, design, and implementation processes of fuzzy databases. This book gives you a first glance at a systematic exposition of the three issues (modeling, design, and implementation). Perhaps the only regret I have in this book is the use of Oracle platform, which, in my view, has the influence of the industrial software of the 1990s. However, the definitions, ideas, and new approaches are platform independent.

Before I say something about the features of the book, I would like to explain some historical aspects that I find interesting to being taken into account by readers. First, in Europe there are two cities well known by the implication of the database in the Zadeh legacy: Toulouse and Granada.

In 1985 Didier Dubois[4] and Henry Prade[5] published *Théorie des Possibilités — Applications à la représentation des connaissances en informatique*, which was translated into English three years later as *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. In Chapter VI of this book the authors introduce the use of the possibility distribution to represent incomplete and uncertain dates in a relational database. This chapter was the result of a PhD thesis written in Toulouse by Claude Testemale[6] and co-directed by Prade. In this work you can see the original code in MACLISP for fuzzy query processing.

Some years later, in Granada, the book of Dubois and Prade, in particular Chapter VI, had a great impact on the PhD thesis of Juan Miguel Medina[7]. In that work Medina summarized the main fuzzy database models in three families (Chapter III): The fuzzy relational model (with a fuzzy degree in each row or tuple), the model based in similarity relations by Buckles and Petry, and the relational models with possibility distributions by Umano, Fukami, Prade, Testemale, Zemankova, Kaendel and other authors. Medina's PhD thesis also embraced the generalizations of fuzzy models. Medina proposed a conceptual framework for fuzzy representation called GEFRED (Generalized Model for Fuzzy Relational Databases) and a language called FSQL (Fuzzy SQL). In the same research group a young mathematician and informatic José Galindo[8] started his PhD research under the supervision of Medina, in order to improve the relational algebra of the GEFRED model, to define a fuzzy relational calculus and to implement other fuzzy comparators. In fact, the possibility and necessity measures, shown by Dubois and Prade, do not only allow the con-

struction of two fuzzy comparators, but 14 of them. The implementation of a new FSQL server running in Oracle and a new GUI interface of the FSQL language was included too.

In these two theses, part of the job was concluded; that is, the physical and logical approaches for development of fuzzy databases. Nevertheless, the conceptual design of fuzzy entities and relations was still missing.

This last step was achieved in 2003 by Angelica Urrutia[9], in her PhD research under the supervision of José Galindo and Mario Piattini. In this work, you find a conceptual fuzzy model, so-called FuzzyEER, and a case tool (FuzzyCASE), to help the database engineers to build the conceptual model for fuzzy databases.

Herein lie the roots of this book, the logical fuzzy models of Medina (1994) and Galindo (1999) on one hand, and, on the other hand, the conceptual fuzzy model of Urrutia (2003).

Personally, I find the name of the book *Fuzzy Databases: Modeling, Design and Implementation* quite right because the work of Galindo, Urrutia, and Piattini is a highly important contribution to understanding the fuzzy database process, not only by professionals of software engineering, but also by computer science students. I hope this book has a real influence in the orientation of the databases courses.

Chapter I, dedicated exclusively to the fuzzy logic, should be appreciated. This chapter could be very useful to new students in this area.

Chapter II brings up to date the classification of fuzzy database models, including some ideas about fuzzy object-oriented database models centered in the relational model, even though these ideas are not used in this book. In spite of this, the contributions of this book will turn out to be very useful for the definition of a complete fuzzy object-oriented database model.

Chapter III is focused on fuzzy database modeling, showing some of the more important approaches by other authors. This chapter is important in order to understanding the importance of the FuzzyEER model defined in Chapter IV, an extension of an EER model to create a model with fuzzy semantics and notations. Although the model has numerous characteristics, the main components of this data modeling tool are: imprecise attributes; fuzzy attributes associated to one or more attributes or with an independent meaning; degrees of fuzzy membership to the model itself, such as fuzzy aggregation, fuzzy entity, weak fuzzy entity, fuzzy relationship; and defined specialization with fuzzy degrees.

Chapter V describes how to represent fuzzy knowledge in relational databases. This methodology is debatable. Nevertheless, as the authors said, it is

complete enough for the immense majority of the applications. On the other hand, the possible lacks in that methodology may be easily solved in each specific application. Chapter VI gives the steps of an algorithm for mapping FuzzyEER models to that methodology. This algorithm relates Chapter IV and V.

Chapter VII describes the more important statements of the FSQL language. This definition improves upon the previous version of this language in many aspects. The educational experience of the authors is noted also in this chapter, which includes a multitude of examples that permits understanding of the utility of each definition.

With all the tools defined in previous chapters, Chapter VIII studies some applications of fuzzy databases. These applications show that fuzzy databases are useful in areas other than management applications (storing and querying information). Of course, FSQL may be used for fuzzy querying, but it can also be used for fuzzy clustering and fuzzy classification, for defining fuzzy dependencies, and for the fuzzy characterization of images in a system of fuzzy image retrieval. The last chapter, the appendices, and references close the book, giving additional information. The open research lines are especially interesting, because they prove that this matter is not closed.

Finally, I borrow the words said by Zadeh in May of 1972, in his Preface of Kaufmann's book, "Professor Kaufmann's treatise is clearly a very important accomplishment. It may a well exert a significant influence on scientific thinking in the years ahead and stimulate much further research on the theory of fuzzy sets and their applications in various field of science and engineering." Well, I think these words match the aim of this book too.

Endnotes

[1]   See http://www.cs.berkeley.edu/~zadeh

[2]   This proposal was mentioned by Zadeh, in the he wrote for the preface of the book written by A. Kaufmann in 1977, *Introduction à la théorie des sous-ensembles flous à l'usage des ingénieurs*.

[3]   The reader can find some ideas related to fuzzy control of engineering systems in the book by Kazuo Tanaka in 1996, *An Introduction to Fuzzy Logic for Practical Applications*.

[4] Dubois, D. (1983). *Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications aux techniques d'aide à la decision.* Doctoral dissertation. Medical and Scientific University, Grenoble, France.

[5] Prade, H. (1982). *Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications au raisonnement naturel.* Doctoral dissertation. Paul Sabatier University, Toulouse, France.

[6] Testemale, C. (1984). *Un système de traitement d'informations incomplètes ou incertaines dans une base de données relationnelles.* Doctoral dissertation. Paul Sabatier University, Toulouse, France.

[7] Medina, J.M. (1994). *Bases de Datos Relacionales Difusas: Modelo teórico y aspecto de su implementación.* Doctoral dissertation. University of Granada, Spain.

[8] Galindo, J. (1999). *Tratamiento de la imprecisión en bases de datos relacionales: Extensión del modelo y adaptación de los SGBD actuales.* Doctoral dissertation. University of Granada, Spain.

[9] Urrutia, A. (2003). *Una definición de un modelo conceptual a una base de datos difusa. Doctoral dissertation.* University of Castilla-La Mancha, Spain.

[10] On February 22, 2005, Dr. Jiménez was awarded with the Trophy Fernand Gallais, who grants the *Ecole Nationale Supérieure des Ingénieurs in Arts Chimiques et Technologiques* of the Polytechnic Institute of Toulouse, France, for his PhD thesis *Gestion des connaissances imparfaites dans les organisations industrielles: Cas d'une industrie manufacturière en Amérique Latine.*

Dr. Leoncio Jiménez[10]
Talca, Chile, June 2005