A Combinatorial Characterization of Fuzzy Functional Dependencies

Anna Czinkóczky-Sali
Department of Information Technology
Bánki Donát Polytechnic
Budapest, P.O.Box 31
H-1428 HUNGARY
E-mail: salias@zeus.banki.hu

Abstract

In the present paper the underlying combinatorial structure of fuzzy functional dependencies is investigated. Equivalence between models of imprecise or uncertain data and a model of changing databases is presented.

1 Introduction

Classical relational database models [8, 13] cannot incorporate, represent or manipulate uncertain or imprecise data. Suppose, a database of teachers of an informatics department is given. How could one represent such statements as "young teacher" or "low/high salary" or "speaks reasonable English"? Clearly, a teacher of age 30 qualifies as young, and one of age 50 does not. But what is the situation with around 40?

In order to be able to handle such issues, fuzziness was introduced. The concept of relational database was extended to include imprecise information. These schemas behave in somewhat similar way as ordinary (crisp) relational schemas. In particular, fuzzy functional dependencies were introduced and several of their properties analysed [2-7]. It turned out that many theorems and concepts valid for (crisp) functional dependencies can be extended to fuzzy functional dependencies. Most notably, fuzzy functional dependencies can be characterized by the fuzzy extension of the Armstrong axioms [8, 13, 4, 5]. The main goal of the present paper is to explore the underlying combinatorial structure. This results in an interesting connection between fuzzy relational databases and another extension of Codd's theory. This latter one is a mathematical model of the changing database.
In the next section the basic definitions of fuzzy relational databases are reviewed. In the third section the main results of the present paper are proved. The fourth section contains the conclusions and some open problems. Preliminary version of this paper was presented in the extended abstracts [9, 10].

2 Fuzzy functional dependencies

In a fuzzy relational database, attribute values are possibility distributions on certain domains with closeness relations [11, 12, 14, 15]. In particular, the following definition was given.

**Definition 1** Let $\Omega = \{A_1, A_2, \ldots, A_n\}$ be a finite set of attributes on the domains $D_1, D_2, \ldots, D_n$, where $\text{domain}(A_i) = D_i$. Let $C = \{c_1, c_2, \ldots, c_n\}$ be a set of closeness relations associated with the domains, i.e., $c_i: D_i \times D_i \mapsto [0, 1]$ where $c_i$ is reflexive ($\forall x \in D_i$, $c_i(x, x) = 1$) and symmetric ($\forall x, y \in D_i$, $c_i(x, y) = c_i(y, x)$). Let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a set of thresholds with $\alpha_i \in [0, 1]$ being specified with respect to $c_i$. Then a relation $R$ of the scheme $R(\Omega, C, A)$ is a subset of $\Pi(D_1) \times \Pi(D_2) \times \ldots \times \Pi(D_n)$ where $\Pi(D_i) = \{\pi_{A_i}: \pi_{A_i} \text{ is an excluding possibility distribution of } A_i \text{ on } D_i\}$. An $n$-tuple $t$ of $R$ is of the form: $t = \{\pi_{A_1}, \pi_{A_2}, \ldots, \pi_{A_n}\}$.

An excluding possibility distribution of $A_i$ on $D_i$ means that the elements of $D_i$ are treated as mutually exclusive. Due to the allowance of fuzzy data to be represented in the database, two values $\pi$ and $\pi'$ of an attribute $A_i$ are not necessarily regarded as equal or not equal, but are regarded as close to each other in a certain degree. A measure for the degree of $\pi = \pi'$ is defined as the possibility that $\pi = \pi'$ is true, based on Zadeh's extension principle assuming normalized possibility distributions:

$$e_c(\pi, \pi') = \sup_{x, y \in D_i, c_i(x, y) \geq \alpha_i} \min(\pi(x), \pi'(y)).$$

This is extended for tuple closeness using min. Let $X = \{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\} \subseteq \Omega$ be a set of attributes and let $t, t' \in R$ be two tuples in the relation. The degree that $t(X)$ equals to $t'(X)$ is defined to be

$$\approx(t(X), t'(X)) = \min(e_c(\pi_{A_{i_1}}, \pi'_{A_{i_1}}), \ldots, e_c(\pi_{A_{i_k}}, \pi'_{A_{i_k}}))$$


**Definition 2** Let $R$ be a relation of the scheme $R(\Omega, C, A)$, $X, Y$ be subsets of $\Omega$, and $I$ be a fuzzy implication operator. Then $X$ functionally determines $Y$ to the degree $\varphi$, $\varphi \in [0, 1]$, denoted by $X \rightarrow_{\varphi} Y$ iff

$$\min_{t, t' \in R} I(\approx(t(X), t'(X)), \approx(t(Y), t'(Y)))) \geq \varphi,$$