

Discovery of Abstract Knowledge from Non-Atomic Attribute Values in Fuzzy Relational Databases

Rafal A. Angryk^a and Frederick E. Petry^{b, c}

^aDepartment of Computer Science, Montana State University
Bozeman, MT 59717-3880, USA
angryk@cs.montana.edu

^bNaval Research Laboratory, Mapping, Charting and Geodesy,
Stennis Space Center, MS 39529, USA
fpetry@nrlssc.navy.mil

^cElectrical Engineering and Computer Science Department, Tulane University
New Orleans, LA 70118, USA

Abstract

In this paper we introduce attribute-oriented induction with partial vote propagation – a new approach allowing acquisition of generalized knowledge from uncertain data. We utilize a proximity-based fuzzy relational database as the medium carrying the original information, where the lack of precise information about an entity is reflected via insertion of multiple attribute values, and the fuzzy relation of α -proximity replaces the classical equivalence relation. Following a well-known approach for generalization of exact data in ordinary databases [9], we introduce a new method for induction of tuples with non-atomic attribute values. In our approach we take advantage of the implicit information about the generalized attributes in the fuzzy database model and apply it to generalize imprecise information.

Keywords: Data Mining, Fuzzy Databases, Attribute-Oriented Induction, Uncertain Data.

1. Introduction

Attribute-Oriented Induction (AOI) is a descriptive database mining technique, which compresses the original set of data into a generalized relation, providing concise and

summarative information about the massive set of the original data. This technique enables a transformation of similar data collections, expressed originally in a database at a low (primitive) level, into more abstract conceptual representations.

The generalization of database records is performed on an attribute-by-attribute basis, utilizing a separate concept hierarchy for each of the generalized attributes included in the relation of task-relevant data. Each concept hierarchy reflects background knowledge about the domain, which is going to be generalized. The concept hierarchies progressively increase the abstraction of the generalization concepts at each new level, allowing the gradual aggregation of attribute values stored in the original tuples.

In contrast to simplified, non-hierarchical data summarization, a gradual attribute-oriented induction through concept hierarchies allows detailed tracking of all records, and can lead to the discovery of interesting patterns among data at the lowest abstraction level of their occurrence. In effect we are able to avoid unnecessary loss of information due to the overgeneralization. Moreover, thorough attribute-oriented induction allows extraction of generalized knowledge, without omitting even rare attribute values. It might occur that such atypical values, despite being initially (at a low level of the generalization hierarchy) infrequent, can sum up to meaningful numbers when generalized to a sufficiently high abstraction level.

In this work we focus on the extraction of generalized knowledge from the imprecise data stored in the fuzzy relational database. In real life, imperfect information occurs very frequently (e.g. caused by the lack of suitable precision during measurements; inconsistency of the data coming from multiple sources; uncertainty in judgments performed by human beings caused by the lack of their objective assessment), and in the miscellaneous areas (e.g. weather maps, genotype characteristics, census data, etc.). An ability to mine knowledge from such data, in spite of the occurring imperfections, has significance for real-world applications. A choice of a suitable approach for generalization of original data can have a fundamental influence on retrieved results, regardless of further data mining techniques.

In the next section we provide a brief review of research conducted on AOI. We also give an overview of a proximity based fuzzy database model, which we utilized as the basis for our generalization approach. Then we present how the fuzzy proximity relation can be applied to build generalization hierarchies, and finally introduce our method allowing attribute-oriented induction of imperfect data.

2 Background

2.1 Attribute-Oriented Induction (AOI)

The idea of applying concept hierarchies to generalize database records for data mining purposes was initially developed by Han et al. [8-10] and extended further by Hamilton et al. [5, 11]. The majority of this work focuses on attribute-oriented induction with utilization of crisp concept hierarchies, where each attribute value (concept) can have only one direct abstract to which it fully belongs.

Several groups of researchers have investigated applications of fuzzy concept hierarchies for AOI. Fuzzy hierarchies of concepts seem to better model real life

dependencies, since they are able to reflect the degree with which one concept belongs to its direct abstract and more than one direct abstract of a single concept is also allowed. Lee and Kim [13] used ISA hierarchies, from area of data modeling, to generalize database records to more abstract concepts. Lee [14] applied fuzzy generalization hierarchies to mine generalized fuzzy quantitative association rules. Cubero et al. [6] presented fuzzy gradual rules for data summarization. Raschia and Mouaddib [17] implemented the SaintEtiq system for data summarization through extended concept hierarchies. Consistent fuzzy concept hierarchies, where each degree of membership is normalized to preserve an exact vote propagation of each tuple when generalized, were recently investigated by the authors [1-2].

2.2 Proximity-based Fuzzy Relational Database

The similarity-based fuzzy model of a relational database, proposed originally by Buckles and Petry [4, 16], is actually a formal generalization of the ordinary relational database model introduced by Codd [7]. The fuzzy model, based on the max-min composition of a fuzzy similarity relation, which replaces the classical equivalence relation coming was further extended by Sheno and Melton [12, 18-19] with the concept of the proximity relation and because of its more general character, we utilized an extension for this approach.

The most distinctive qualities of the fuzzy relational database are: (1) allowing non-atomic domain values, when characterizing attributes of an entity and (2) generation of equivalence classes with the support of a proximity relation applied in the place of traditional identity relation.

As mentioned above, each attribute value of the fuzzy database record is allowed to be a subset of the whole base set of attribute values describing a particular domain. Formally, if we denote a set of acceptable attribute values as D_j , and we let d_{ij} to symbolize a particular (i^{th}) attribute value, characterizing the i^{th} entity; the original Codd's principle $d_{ij} \in D_j$ is replaced in the fuzzy database schema with the set $d_{ij} \subseteq D_j$. Any member of the power set of accepted domain values can be inserted as an attribute descriptor except the null set. Therefore a fuzzy database relation is actually a subset of the cross product of all power sets of its constituent attributes $2^{D_1} \times 2^{D_2} \times \dots \times 2^{D_n}$. This property permits fuzzy database to store imprecision coming from the original source of information. In cases when the particular entity cannot be clearly characterized by a single descriptor, the uncertainty can be reflected by multiple attribute values.

A proximity-based fuzzy database uses an explicitly declared proximity relation of which both the identity and similarity relations are special cases. Since a fuzzy proximity relation (also called tolerance relation) is only reflexive and symmetric, which is not sufficient to obtain equivalence classes, transitivity of proximity relation was an extension. This is achieved by modifying the original definition of fuzzy proximity relations with transitivity via similarity paths (sequences of similarities), using Tamura chains [20]. The definition of α -proximity relation employed in this approach is:

If P is a proximity relation on D_j , then given an $\alpha \in [0, 1]$, two elements $x, z \in D_j$ are α -similar (denoted by $x P_{\alpha} z$) if and only if $P(x, z) \geq \alpha$, and are said to be α -proximate (denoted

by $x P_a^+ z$) if and only if they are (1) either *a-similar* or (2) there exists a sequence $y_1, y_2, \dots, y_m \hat{I} D_j$ such that $x P_a y_1 P_a y_2 P_a \dots P_a y_m P_a z$.

Table 1: Proximity table for a domain HAIR COLOR.

	<i>black</i>	<i>d. brown</i>	<i>auburn</i>	<i>red</i>	<i>blond</i>	<i>bleached</i>
<i>black</i>	1.0	0.8	0.6	0.5	0.3	0.1
<i>d. brown</i>	0.8	1.0	0.7	0.6	0.5	0.2
<i>auburn</i>	0.6	0.7	1.0	0.8	0.4	0.3
<i>red</i>	0.5	0.6	0.8	1.0	0.5	0.4
<i>blond</i>	0.3	0.5	0.4	0.5	1.0	0.8
<i>bleached</i>	0.1	0.2	0.3	0.4	0.8	1.0

Each of the attributes in the fuzzy database has its own *proximity table*, which includes the *degrees of proximity* (called above *a-similarity*) between all values occurring for the particular attribute. A proximity table for the domain of HAIR COLOR, which we will use as an example for our further analysis, is presented in the Table 1. Since it is not necessary to preserve max-min transitivity when defining the proximity values the relation becomes much easier to specify.

As proposed by Tamura [20], the proximity table can be transformed to reflect the *a-proximity relation*. The results of such a transformation are seen in Table 2.

Table 2: α -proximity table for a domain HAIR COLOR.

	<i>black</i>	<i>d. brown</i>	<i>auburn</i>	<i>red</i>	<i>blond</i>	<i>bleached</i>
<i>black</i>	1.0	0.8	0.7	0.7	0.5	0.5
<i>d. brown</i>	0.8	1.0	0.7	0.7	0.5	0.5
<i>auburn</i>	0.7	0.7	1.0	0.8	0.5	0.5
<i>red</i>	0.7	0.7	0.8	1.0	0.5	0.5
<i>blond</i>	0.5	0.5	0.5	0.5	1.0	0.8
<i>bleached</i>	0.5	0.5	0.5	0.5	0.8	1.0

Now the disjoint classes of attribute values, considered to be equivalent at the specific α -level, can be extracted from the table. They are marked by shadings in Table 2. Such a separation of the equivalence classes arises due to the sequential similarity proposed by Tamura. For instance, despite the fact that the proximity degree, presented in Table 1, between the concepts *red* and *black* is 0.5, the α -proximity is 0.7. Via the sequence of the original proximity degrees: *black* P_a *d.brown*=0.8 \hat{U} *d.brown* P_a *auburn*=0.7 \hat{U} *auburn* P_a *red*=0.8, we get *black* P_a^+ *red*=0.7, which is shown in Table 2.

3. Attribute-Oriented Induction from fuzzy tuples

3.1 Building concept hierarchy from a α -proximity table

The creation of an α -proximity relation for a particular domain D_j can lead to the extraction of a crisp concept hierarchy, allowing attribute-oriented induction on such a domain. From the propagation of shadings in the Table 2, we can easily observe that the equivalence classes marked in the table have a nested character.

As in the case of a fuzzy similarity relation [21], each α -cut (where $\alpha \in [0, 1]$) of a fuzzy binary relation in Table 2 creates disjoint equivalence classes in the domain D_j . If we let P_α denote a single equivalence class partition induced on domain D_j by a single α -level-set, then by an increase of the value of α to α' we are able to extract the subclass of P_α , denoted $P_{\alpha'}$ (a refinement of the previous equivalence class partition). A nested sequence of partitions $P_{\alpha^1}, P_{\alpha^2}, \dots, P_{\alpha^k}$, where $\alpha^1 < \alpha^2 < \dots < \alpha^k$, may be represented in the form of a *partition tree*, as in Figure 1.

This nested sequence of partitions in the form of a tree has a structure identical with the crisp concept hierarchy used for AOI. The increase of conceptual abstraction in the partition tree is denoted by decreasing values of α ; lack of abstraction during generalization (0-abstraction level at the bottom of generalization hierarchy) complies with the 1-cut of the similarity relation ($\alpha=1.0$), and is denoted as $S_{1,\alpha}$.

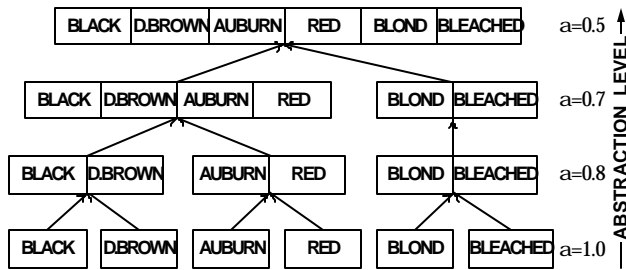


Figure 1: Partition tree of domain HAIR COLOR, built on the basis of Table 2.

An advantage of the utilization of the proximity-based fuzzy model is that such a hierarchy, by definition implemented in every such fuzzy database, can be extracted automatically for a user who has no background knowledge about the specific domain.

The only thing differentiating the hierarchy in the Figure 1 from the crisp concept hierarchies applicable for AOI is the lack of abstract concepts, which are used as the labels characterizing the sets of generalized (grouped) concepts. To create a complete set of the abstract labels it is sufficient to choose only one value of the attribute per the equivalence class at the each level of hierarchy (α), and assign a unique abstract descriptor to it. Sets of such definitions (value of attribute and value of α linked with an

abstract name) can be stored as a relational database table (Table 3), where the first two columns create a natural key for this relation.

Table 3: Table of abstract descriptors (for Figure1).

<i>ATTRIBUTE VALUE</i>	<i>ABSTRACTION LEVEL (a)</i>	<i>ABSTRACT DESCRIPTOR</i>
<i>black</i>	<i>0.8</i>	<i>DARKISH</i>
<i>red</i>	<i>0.8</i>	<i>REDDISH</i>
<i>blond</i>	<i>0.8</i>	<i>BLONDISH</i>
<i>black</i>	<i>0.7</i>	<i>DARK</i>
<i>blond</i>	<i>0.7</i>	<i>BLONDISH</i>
<i>black</i>	<i>0.5</i>	<i>ANY</i>

The combination of partition tree in Figure 1 and the table of abstract descriptors allow us to build the generalization hierarchy in the form shown in Figure 2.

The disjoint character of equivalence classes generated from the α -proximity table does not allow any concept in the hierarchy to have more than one direct abstract at every level of generalization hierarchy. Therefore this approach can be utilized only to form a crisp generalization hierarchy. Such a hierarchy, however, can be then successfully applied as a foundation to the development of a fuzzy concept hierarchy – by extending it with additional edges to represent partial membership of the lower level concepts in their direct abstract descriptors. Depending on the assigned memberships, reflecting preferences of the user, this can create consistent or inconsistent fuzzy concept hierarchies.

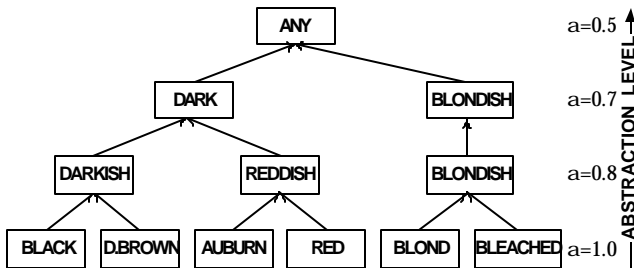


Figure 2: Crisp generalization hierarchy formed using Tables 2 and 3.

3.2 Character of imprecision reflected in fuzzy records

Before introducing our approach to AOI from imprecise data, let us analyze briefly the nature of the uncertainty representation allowed in the fuzzy database model. There are two actual representations of imprecision in the fuzzy database schema. First, as already

mentioned, is the occurrence of multiple attribute values. Obviously, the more descriptors we use to characterize a particular record in the database, the more imprecise is its depiction. Uncertainty about the description is also implicitly reflected in the similarity of values characterizing a particular entity. e.g. when we describe someone's hair as *{black, dark brown, red, auburn}* we have more doubt about the person's hair colour than in the case when we characterize it as *{blond, dark blond, light brown, brown}*, since this description would be rather immediately interpreted as "blondish". There are the same number of attribute values in each case, however the higher similarity of values utilized in the second set results in the higher informativeness carried by the second example.

The imprecision of the original information is actually reflected both in the number of inserted descriptors for a particular attribute and in the similarity of these values. In Table 4 we summarize observations concerning their relationship. The domain called *Quantity of attribute values* is a discrete set of integer numbers (> 0 , since the fuzzy model does not allow empty attributes); the *Similarity of attribute values* is characterized in fuzzy databases with a continuous set of real numbers in a range $[0, 1]$ – the values of α .

Table 4: Character of information stored in the Fuzzy Databases.

Quantity of attr. values \ Similarity of attr. values	LOW	HIGH
	SMALL	<i>Imprecise</i>
LARGE	<i>Imprecise</i> <i>(Error suspected)</i>	<i>Precise</i> <i>(Confirmed)</i>

The simplified characterization of data imprecision presented in Table 4 can be enhanced with a brief analysis of the boundary values. The measure of imprecision can be thought of ranging between 0 (i.e. the lack of uncertainty about results) and infinity (maximum imprecision). The common opinion that even flawed information is better than lack of the information, leads us to say that imprecision reaches its maximum limits when there is no data inserted at all. Since the fuzzy database model does not allow empty attributes we will not consider this further. The minimum imprecision (0-level) is achieved by a single attribute value. If there are no other descriptors or auxiliary information, we must assume the inserted value is a perfect characterization of the particular entity's feature. The same minimum can be also accomplished with multiple values if they all have identical meaning (synonyms). Despite the fact that multiple, identical descriptors additionally confirm an initially inserted value, they cannot lead to further reduction of imprecision, since it already has the minimal value. Therefore the descriptors, which are so similar that they are considered to be identical, can be reduced to a single descriptor. Obviously, some attribute values, initially considered as different, may be treated as identical at a higher abstraction level. Therefore we can conclude that the

practically achievable minimum of imprecision depends on the abstraction level of employed descriptors, and can reach its original θ level only at the lowest level of abstraction (for $\alpha=1.0$ in our fuzzy database model).

3.3 Partial Vote Propagation to generalize imprecise data

Since the fuzzy database model permits the reflection of uncertainty about the value characterizing each feature via insertion of multiple attribute descriptors, it is necessary to provide a mechanism allowing AOI from such data. In this section we propose a method enabling generalization of multiple attribute values, based on the dependencies presented in the previous section.

In the case of attribute generalization with utilization of a concept hierarchy (Figure 2), we have single attribute values at the bottom level of the hierarchy. Therefore the generalization of tuples with single descriptors is straightforward. A problem arises with the case of multiple attribute values describing a single entity. Where should we assign a person whose hair was described as $\{d.brown, auburn, red\}$? Our solution is based on partial vote propagation, where a single vote, corresponding to one database record, is partitioned to represent each of the originally inserted attribute values. During AOI all fractions of this vote propagate gradually through multiple levels of generalization hierarchy, the same way as the regular (precise) records do. The only difference is that the record with uncertainty has multiple generalization paths (different paths for different vote's fractions), whereas each of the precise records has only one generalization path.

The most trivial solution would be to split the vote equally among all inserted descriptors: $\{d.brown/0.(3), auburn/0.(3), red/0.(3)\}$. This approach however does not take into consideration real life dependencies, which are reflected not only in the number of inserted descriptors, but also in their similarity. We propose replacement of the even distribution of vote with a nonlinear spread, dependent on the similarity and the number of inserted values. Using the partition tree (Figure 1), we can extract from the set of the originally inserted values the concepts which are more similar to each other than to the remaining descriptors; we call these subsets of resemblances (e.g. $\{red, auburn\}$ from the above-mentioned example). Then we use the subset as a base for calculating new vote's fractions. An important aspect of this approach is extraction of the subsets of similarities at the lowest possible level of their occurrence, since the nested character of α -proximity relation guarantees that above this α -level they will always occur together. Repetitive extraction of such subsets could unbalance the original dependencies among inserted values.

The proposed approach is rather straightforward given (1) a set of attribute values inserted as a description of particular entity (i.e. *Set of Descriptors*), and (2) a hierarchical structure (tree) reflecting Zadeh's partition tree for the particular attribute (Figure 1). We want to extract the list of all subsets of similarities from the given *Set of Descriptors*, with the highest *Level of α -proximity* of their common occurrence. This is achieved by preorder recursive traversal of the partition tree.. Searching from the root of the tree, if any subset of the given *Set of Descriptors* occurs at a particular node of the concept hierarchy, we store the values that were recognized as similar, and the value of α . In Figure 3 we present an example of such a search for subsets of similarities for a record

with values $\{black, d.brown, blond, red\}$. Numbers on the links in the tree represent the order in which the particular subsets of similarities were extracted.

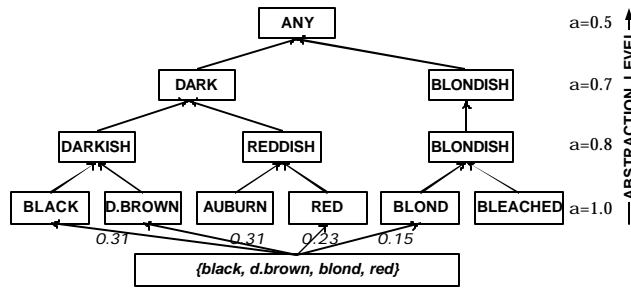


Figure 3: Subsets of similar values extracted from the original set of descriptors.

After extracting the subsets of similarities (Figure 3), we summarize α values as a measure reflecting both the frequency of occurrence of the particular attribute values in the subsets of similarities, as well as the abstraction level of these occurrences. Since value *blond* appeared only at the top and the bottom level, we assign it a grade 1.5 ($1.0+0.5$). The remaining attribute values were graded as follows:

$$black/(1.0 + 0.8 + 0.7 + 0.5) = black/3.0$$

$$d.brown/(1.0 + 0.8 + 0.7 + 0.5) = d.brown/3.0$$

$$red/(1.0 + 0.7 + 0.5) = red/2.2$$

In the next step, we use the sum of all generated grades ($1.5+3.0 + 3.0 + 2.2 = 9.7$) in order to normalize the grades finally assigned to each of the participating attribute values:

$$black/(3.0/9.7) = black/0.31$$

$$d.brown/(3.0/9.7) = d.brown/0.31$$

$$red/(2.2/9.7) = red/0.23$$

$$blond/(1.5/9.7) = blond/0.15$$

This new distribution of the vote's fractions more accurately reflects real life dependencies than a linear approach. The final results are shown in Figure 4.

Normalization of the initial grades has a crucial meaning for preservation of the generalization model's completeness. It guarantees that each of the records is represented as a unity, despite being variously distributed at each of the generalization levels.

During the AOI process all fractions of the vote may gradually merge to finally become unity at the level of abstraction high enough to erase the originally occurring imprecision. In such a case, we observe that there is a removal of imprecision from data due to its generalization. Such a connection between the precision and certainty seems to be natural and was already noted by other researchers [3, 15]. In general, very abstract statements have a greater probability to be "correct" than more detailed ones.

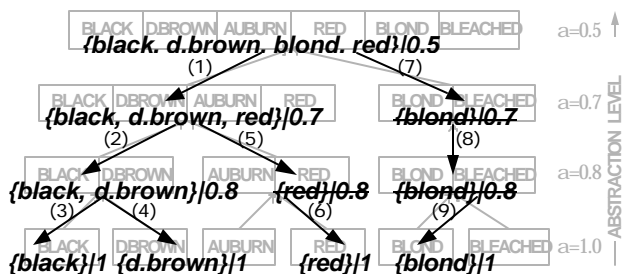


Figure 4: Partial Vote Propagation for records with uncertainty.

4 Conclusions

An acceptance of non-atomic values for an attribute may lead to the occurrence of imprecision, when a single attribute is described by multiple values which are not considered to be equal at the given level of detail. In this paper we presented a new approach allowing Attribute-Oriented Induction in such cases.

In our approach, multiple descriptors with high similarity are allowed to dominate distribution of the vote's fractions, which seems to more accurately reflect the actually occurring dependencies. Such a nonlinear distribution of the record's vote, based on the similarities among attribute values, implicitly supports a reduction of the original data's inconsistency.

In this paper we utilized a proximity-based fuzzy relational database model, however the proposed mechanism of generalization via partial vote distribution based on the shape of concept hierarchy seems to be applicable to AOI on any data with non-atomic attribute values. Applying this approach to mine abstract knowledge from imprecise data stored in other types of databases remains an interesting challenge for our future work.

Acknowledgements

We would like to thank the Naval Research Laboratory's Base Program, Program Element No. 0602435N for sponsoring this research.

References

1. R.A. Angryk and F.E. Petry, "Consistent fuzzy concept hierarchies for attribute generalization," *Proceeding of the IASTED International Conference on Information and Knowledge Sharing (IKS '03)*, Scottsdale, AZ, USA, November 2003, pp. 158-163.
2. R.A. Angryk and F.E. Petry, "Data Mining Fuzzy Databases Using Attribute-Oriented Generalization," *Proceeding of The Third IEEE International Conference on Data Mining (ICDM '03), Workshop on Foundations and New Direction in DM*, Melbourne, FL, USA, November 2003, pp. 8-15.

3. P. Bosc and H. Prade, "An introduction to fuzzy set and possibility theory based approaches to the treatment of uncertainty and imprecision in database management systems", *Proceedings of the 2nd Workshop on Uncertainty Management in Information Systems (UMIS'94): From Needs to Solutions*, Catalina, CA, USA, 1994.
4. B.P. Buckles and F.E. Petry, "A fuzzy representation of data for relational databases", *Fuzzy Sets and Systems*, 7(3), 1982, pp. 213-226.
5. C.L. Carter and H.J. Hamilton, "Efficient Attribute-Oriented Generalization for Knowledge Discovery from Large Databases", *IEEE Transactions on Knowledge and Data Engineering*, 10(2), 1998, pp. 193-208.
6. J. C. Cubero, J.M. Medina, O. Pons, M.A. Vila, "Data Summarization in Relational Databases through Fuzzy Dependencies", *Inf. Sciences*, 121(3-4), 1999, pp. 233-270.
7. F.E. Codd, "A relational model of data for large share data banks", *Communications of the ACM*, 13(6), 1970, pp. 377-387.
8. J. Han, Y. Cai, and N. Cercone, "Knowledge discovery in databases: An attribute-oriented approach", *Proc. 18th Int. Conf. VLDB*, Vancouver, Canada, 1992, pp. 547-559.
9. J. Han and Y. Fu, "Exploration of the Power of Attribute-Oriented Induction in Data Mining", in U.M. Fayyad, G. Hatetsky-Shapiro, P. Smyth, and R. Uthurusamy (Eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, Menlo Park, CA, 1996, pp. 399-421.
10. J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, New York, NY, 2000.
11. R.J. Hilderman, H.J. Hamilton, and N. Cercone, "Data mining in large databases using domain generalization graphs", *Journal of Intelligent Information Systems*, 13(3), 1999, pp. 195-234.
12. S. Kumar De, R. Biswas, and A. R. Roy, "On extended fuzzy relational database model with proximity relations", *Fuzzy Sets and Systems*, 117(2), 2001, pp.195-201.
13. D.H. Lee and M.H. Kim, "Database summarization using fuzzy ISA hierarchies", *IEEE Transactions on Systems, Man, and Cybernetics - part B*, 27(1), 1997, pp. 68-78.
14. K.-M. Lee, "Mining generalized fuzzy quantitative association rules with fuzzy generalization hierarchies", *Proc. Joint 9th IFSA World Congress and 20th NAFIPS Int'l Conf.*, Vancouver, Canada, 2001, pp. 2977-2982.
15. S. Parsons, Current approaches to handling imperfect information in data and knowledge bases, *IEEE Trans. on Knowledge and Data Eng.*, 8(3), 1996, pp. 353-372.
16. F.E. Petry, *Fuzzy Databases: Principles and Applications*, Kluwer Academic Publishers, Boston, MA, 1996.
17. G. Raschia and N. Mouaddib, "SAINTETIQ: a fuzzy set-based approach to database summarization", *Fuzzy Sets and Systems*, 129(2), 2002, pp. 137 -162.
18. S. Shenoï and A. Melton, "Proximity Relations in the Fuzzy Relational Database Model", *International Journal of Fuzzy Sets and Systems*, 3(3), 1989, pp. 285-296.
19. S. Shenoï, A. Melton, and L. T. Fan, "Functional Dependencies and Normal Forms in the Fuzzy Relational Database Model", *Information Sciences*, 60(1-2), 1992, pp. 1-28.
20. S. Tamura, S. Higuchi, and K. Tanaka, "Pattern Classification Based on Fuzzy Relations", *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-1(1), 1971, pp. 61-66.
21. L.A. Zadeh, "Similarity relations and fuzzy orderings", *Information Sciences*, 3(2), 1970, pp. 177-200.