A Fuzzy FCA-based Approach to Conceptual Clustering for Automatic Generation of Concept Hierarchy on Uncertainty Data

Thanh Tho Quan[1], Siu Cheung Hui[1], and Tru Hoang Cao[2]

[1]School of Computer Engineering, Nanyang Technological University, Singapore
{PA0218164B, asschui}@ntu.edu.sg
[2]Faculty of Information Technology, Hochiminh City University of Technology, Vietnam
tru@dit.hcmut.edu.vn

Abstract. This paper proposes a new fuzzy FCA-based approach to conceptual clustering for automatic generation of concept hierarchy on uncertainty data. The proposed approach first incorporates fuzzy logic into Formal Concept Analysis (FCA) to form a fuzzy concept lattice. Next, a fuzzy conceptual clustering technique is proposed to cluster the fuzzy concept lattice into conceptual clusters. Then, hierarchical relations are generated among conceptual clusters for constructing the concept hierarchy. In this paper, we also apply the proposed approach to generate a concept hierarchy of research areas from a citation database. The performance of the proposed approach is also discussed in the paper.

Keywords: Formal Concept Analysis, Fuzzy Logic, Conceptual Clustering, Concept Hierarchy

1 Introduction

Conceptual clustering [3] is an advanced data mining technique that clusters data into clusters associated with conceptual representations, or *conceptual clusters*. Concept hierarchy can then be constructed from the conceptual clusters. However, traditional conceptual clustering techniques can only work on specific data types such as nominal and numeric. In addition, the concept hierarchy is mostly in a tree-like structure which is unable to support the representation of multiple inheritance.

Formal Concept Analysis (FCA) [4] is a data analysis technique based on the ordered lattice theory. It defines formal contexts to represent relationships between objects and attributes in a domain and interprets the corresponding concept lattice. The concept lattice is more informative than traditional tree-like conceptual structures as it can also support multiple inheritance. This makes FCA a very suitable technique for conceptual clustering. Several FCA-based

© V. Snášel, R. Bělohlávek (Eds.): CLA 2004, pp. 1–12, ISBN 80-248-0597-9.
VŠB – Technical University of Ostrava, Dept. of Computer Science, 2004.

conceptual clustering systems such as TOSCANA [15] and INCOSHAM [6] have been developed.

However, there are many situations in which uncertainty information also occurs. For example, keywords extracted from scientific documents can be used to infer the corresponding research areas, however, it is inappropriate to treat all keywords equally as some keywords may be more significant than others. Moreover, it is sometimes difficult to judge whether a document belongs totally to a research area or not. Traditional FCA-based conceptual clustering approaches are hardly able to represent such vague information. To tackle this problem, we propose a fuzzy FCA-based approach to conceptual clustering for automatic generation of concept hierarchy on uncertainty data.

Pollandt [13], and Huynh and Nakamori [7] proposed the L-Fuzzy context as an attempt to combine fuzzy logic with FCA. The L-Fuzzy context uses linguistic variables, which are linguistic terms associated with fuzzy sets, to represent uncertainty in the context. However, human interpretation is required to define the linguistic variables. Moreover, the fuzzy concept lattice generated from the L-fuzzy context usually causes a combinatorial explosion of concepts as compared to the traditional concept lattice.

In this paper, we propose a new technique that incorporates fuzzy logic into FCA as Fuzzy Formal Concept Analysis (FFCA), in which uncertainty information is directly represented by a real number of membership value in the range of [0,1]. As such, linguistic variables are no longer needed. In comparison with the fuzzy concept lattice generated from the L-fuzzy context, the fuzzy concept lattice generated using FFCA will be simpler in terms of the number of formal concepts, and it also supports a formal mechanism for calculating concept similarities. Therefore, the proposed FFCA's fuzzy concept lattice is a suitable representation for conceptual clustering.

The rest of the paper is organized as follows. Section 2 discusses the related work on conceptual clustering. Section 3 presents the proposed approach. Section 4 discuses the Fuzzy Formal Concept Analysis. Fuzzy Conceptual Clustering is presented in Section 5. Section 6 discusses the Hierarchical Relation Generation process. Section 7 applies the proposed approach to a citation database for generating a concept hierarchy of research areas. Section 8 evaluates the performance of the proposed approach. Finally, Section 9 concludes the paper.

2 Conceptual Clustering

Conceptual clustering techniques can be used to construct a concept hierarchy from data. Traditional conceptual clustering techniques such as COBWEB [3] and AutoClass [1] are based on taxonomy clustering techniques and use statistical models as conceptual representations of clusters. However, these techniques are only applicable to specific types of data. CLASSIT [5] and ECOBWEB [14] were proposed to improve on COBWEB to deal with numeric attributes of data. SBAC [10] was introduced as a conceptual clustering technique that can handle mixed numeric and nominal data. However, as conceptual hierarchies generated by these techniques are represented as tree-like structures, multiple inheritance is not supported. In this paper, we propose a fuzzy FCA-based approach for conceptual clustering that can handle uncertainty data and represent the data in a fuzzy concept lattice.

3 Proposed Approach

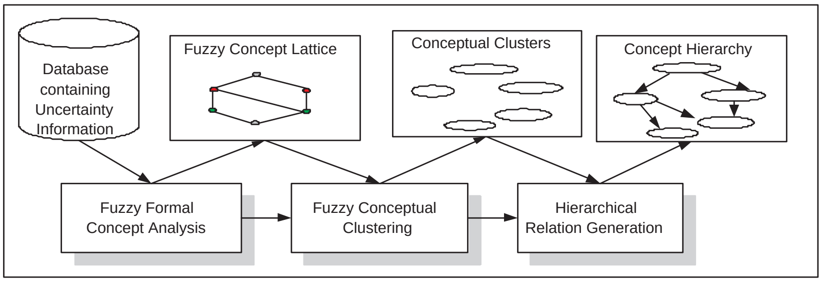

Fig. 1. The proposed approach for automatic generation of concept hierarchy

Figure 1 shows the proposed fuzzy FCA-based approach to conceptual clustering for automatic generation of concept hierarchy. It consists of the following steps: Fuzzy Formal Concept Analysis,Fuzzy Conceptual Clustering and Hierarchical Relation Generation.

4 Fuzzy Formal Concept Analysis

In this section, we discuss the Fuzzy Formal Concept Analysis, which incorporates fuzzy logic into Formal Concept Analysis, to represent vague information.

Definition 1. *A fuzzy formal context is a triple* $K = (G, M, I = \varphi(G \times M))$ *where* G *is a set of objects,* M *is a set of attributes, and* I *is a fuzzy set on domain* $G \times M$*. Each relation* $(g, m) \in I$ *has a membership value* $\mu(g, m)$ *in* $[0, 1]$.

A fuzzy formal context can also be represented as a cross-table as shown in Table 1(a). The context has three objects representing three documents, namely $D1$, $D2$ and $D3$. In addition, it also has three attributes, "**D**ata Mining" (D), "**C**lustering" (C) and "**F**uzzy Logic" (F) representing three research topics. The relationship between an object and an attribute is represented by a membership value between 0 and 1.

A *confidence threshold* T can be set to eliminate relations that have low membership values. Table 1(b) shows the cross-table of the fuzzy formal context given in Table 1(a) with $T = 0.5$.

Table 1(a). A cross-table of a fuzzy formal context.

	D	**C**	**F**
D1	0.8	0.12	0.61
D2	0.9	0.85	0.13
D3	0.1	0.14	0.87

Table 1(b). Fuzzy formal context in Table 1(a) with T =0.5.

	D	**C**	**F**
D1	0.8	-	0.61
D2	0.9	0.85	-
D3	-	-	0.87

Generally, we can consider the attributes of a formal concept as the description of the concept. Thus, the relationships between the object and the concept should be the intersection of the relationships between the objects and the attributes of the concept. Since each relationship between the object and an attribute is represented as a membership value in fuzzy formal context, then the intersection of these membership values should be the minimum of these membership values, according to fuzzy theory [16].

Definition 2. *Given a fuzzy formal context* K=(G, M, I) *and a confidence threshold* T, *we define* $A*= \{m \in M | \forall g \in A: \mu(g, m) \geq T\}$ *for* $A \subseteq G$ *and* $B*= \{g \in G | \forall m \in B: \mu(g,m) \geq T\}$ *for* $B \subseteq M$. *A fuzzy formal concept (or fuzzy concept) of a fuzzy formal context* (G, M, I) *with a confidence threshold* T *is a pair* (A_f =$\varphi(A)$, B) *where* $A \subseteq G$, $B \subseteq M$, $A* = B$ *and* $B* = A$. *Each object* $g \in \varphi(A)$ *has a membership* μ_g *defined as*

$$\mu_g = \min_{m \in B} \mu(g, m)$$

where $\mu(g, m)$ *is the membership value between object* g *and attribute* m, *which is defined in* I. *Note that if* $B = \{\}$ *then* μ_g = *1 for every* g.

Definition 3. *Let* (A_1, B_1) *and* (A_2, B_2) *be two fuzzy concepts of a fuzzy formal context* (G, M, I). ($\varphi(A_1)$, B_1) *is the subconcept of* ($\varphi(A_2)$, B_2), *denoted as* ($\varphi(A_1)$, B_1) $\leq$ ($\varphi(A_2)$, B_2), *if and only if* $\varphi(A_1) \subseteq \varphi(A_2) (\Leftrightarrow B_2 \subseteq B_1)$. *Equivalently,* ($A_2$, B_2) *is the superconcept of* (A_1, B_1).

Definition 4. *A fuzzy concept lattice of a fuzzy formal context* K *with a confidence threshold* T *is a set* F (K) *of all fuzzy concepts of* K *with the partial order* $\leq$ *with the confidence threshold* T.

Definition 5. *The similarity of a fuzzy formal concept* K_1= ($\varphi(A_1)$, B_1) *and its subconcept* K_2 = ($\varphi(A_2)$, B_2) *is defined as* $E(K_1, K_2) = \frac{|\varphi(A_1) \cap \varphi(A_2)|}{|\varphi(A_1) \cup \varphi(A_2)|}$, *where* $\cap$ *and* $\cup$ *refer intersection and union operators on fuzzy sets, respectively.*

Figure 2 gives the traditional concept lattice generated from Table 1(a), in which crisp values "Yes" and "No" are used instead of membership values. Figure 3 gives the fuzzy concept lattice generated from the fuzzy formal context given in Table 1(b). As shown from the figures, the fuzzy concept lattice can provide additional information, such as membership values of objects in each fuzzy formal concept and similarities of fuzzy formal concepts, that are important for the construction of concept hierarchy.

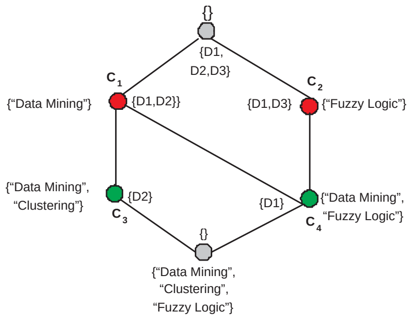

Fig. 2. A concept lattice generated from traditional FCA.

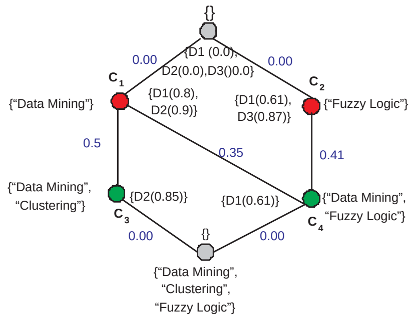

Fig. 3. A fuzzy concept lattice generated from FFCA.

5 Fuzzy Conceptual Clustering

As in traditional concept lattice, the fuzzy concept lattice generated using FFCA is sometimes quite complicated due to the large number of fuzzy formal concepts generated. Since the formal concepts are generated mathematically, objects that have small differences in terms of attribute values are classified into distinct formal concepts. At a higher level, such objects should belong to the same concept when they are interpreted by human.

Based on this observation, we propose to cluster formal concepts into conceptual clusters using fuzzy conceptual clustering, which have the following properties:

- Each conceptual cluster is a sublattice extracted from the fuzzy concept lattice.
- A formal concept must belong to at least one conceptual cluster, but it can also belong to more than one conceptual cluster. This property is derived from the characteristic of concepts that an object can belong to more than one concept. For example, a scientific document can belong to more than one research area.

Conceptual clusters are generated based on the premise that if a formal concept A belongs to a conceptual cluster R, then its subconcept Balso belongs to R if B is similar to A. We can use a *similarity confidence threshold* T_s to determine whether two concepts are similar or not.

Definition 6. *A conceptual cluster of a concept lattice K with a similarity confidence threshold T_s is a sublattice S_K of K which has the following properties:*

1. *S_K has a supremum concept C_S that is not similar to any of its superconcepts.*
2. *Any concept $C \neq C_S$ in S_K must have at least one superconcept $C' \in S_K$ such that $E(C, C') > T_s$.*

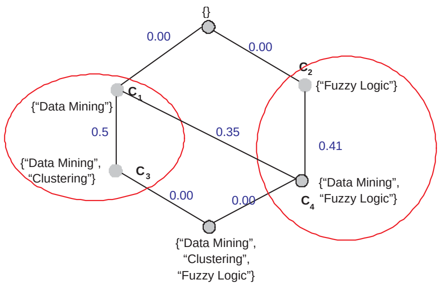

Fig. 4. Conceptual clusters generated from Figure 2(b) with confidence threshold Ts = 0.4.

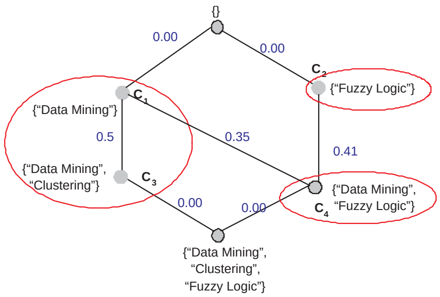

Fig. 5. Conceptual clusters generated from Figure 2(b) with confidence threshold Ts = 0.5.

Figures 4 and 5 show the conceptual clusters that are generated from the fuzzy concept lattice given in Figure 3 with the similarity confidence thresholds $T_s = 0.4$ and $T_s = 0.5$ respectively.

Algorithm: Conceptual_Cluster_Generation

Input: Starting concept C_S of concept lattice $F(K)$ and a similarity threshold T_S

Output: A set of generated conceptual clusters S_C

Process:

$S_C \leftarrow \{\}$

$F'(K) \leftarrow$ An empty concept lattice

Add C_S to $F'(K)$

for each subconcept C' of $C_S$in $F(K)$ **do**

  $F'(C') \leftarrow$ Conceptual_Cluster_Generation(C', $F(K)$, T_S)

  if $E(C_S, C') = \frac{|C_S \cap C'|}{|C_S \cup C'|} < T_S$ **then**

    $S_C \leftarrow S_C \cup \{ F'(C')\}$

  else

    Insert $F'(C')$ to $F'(K)$ with sup($F'(K)$) as a subconcept of C_S

  endif

endfor

$S_C \leftarrow S_C \cup \{F'(K)\}$

Fig. 6. The fuzzy conceptual clustering algorithm.

Figure 6 gives the algorithm that generates conceptual clusters from a concept C_S which is called the *starting concept* on a fuzzy concept lattice $F\ (K)$. To generate all conceptual clusters of $F\ (K)$, we choose C_S as the supremum of $F(K)$, or C_S =sup $(F(K))$.

6 Hierarchical Relation Generation

As discussed in Section 5, fuzzy conceptual clustering generates a set of conceptual clusters S_C. To construct a concept hierarchy from the conceptual clusters, we need to find the hierarchy relations from the clusters. We first define a concept hierarchy [11] as follows.

Definition 7. *A concept hierarchy is a poset (partially ordered set)* $(H,\angle)$ *where* H *is a finite set of concepts, and* $\angle$ *is a partial order on* H.

Hence, to construct a concept hierarchy satisfying Definition 7, we must construct hierarchical relations among conceptual clusters as a partial order on S_C.

Definition 8. *Let* C_1 *and* C_2 *be two conceptual clusters corresponding to two sublattices* L_1 *and* L_2 *of a fuzzy concept lattice* $F(K)$. *Let the fuzzy formal concept* I *be the supremum of* L_1, *or* $I = sup(L_1)$. C_1 *is the subconcept of* C_2, *denoted as* $C_1 \angle C_2$, *if* I *is the subconcept of any concept* $C' \in L_2$, *, or* $I \leq C'$ *where* $\leq$ *is the partial order defined on* $F(K)$. *Equivalently,* C_2 *is the superconcept of* C_1.

From the FCA theory, a fuzzy concept lattice $F(K)$ is a complete lattice. That is, any fuzzy concept C_F on $F(K)$ must have at least one superconcept unless $C_F = \sup(F(K))$. Therefore, the definition of superconcept and subconcept relations on conceptual clusters, which is given in Definition 8, assures that each conceptual cluster has at least one superconcept, unless it corresponds to the root node of the concept hierarchy generated. However, we must prove that the $\angle$ relation given in Definition 8 is a partial order.

Corollary 1. *Let* C_1 *and* C_2 *be two conceptual clusters corresponding to two sublattices* L_1 *and* L_2 *of a fuzzy concept lattice* $F(K)$. *Let the fuzzy formal concepts* I_1 *and* I_2 *be the supremums of* L_1 *and* L_2 *respectively. If* $C_1 \angle C_2$ *then* $I_1 \leq I_2$ *where* $\leq$ *is the partial order defined on* $F(K)$.

Proof. Since $C_1 \angle C_2$, then $\exists C' \in L_2$ such that $I_1 \leq C'$. Since $I_2 = \sup\ (L_2)$, therefore $C' \leq I_2$. Since $\leq$ is a partial order, then $I_1 \leq I_2$.□

From Corollary 1, we realize that the $\angle$ relation of conceptual clusters is equivalent to the $\leq$ relation of the supremums of the corresponding sublattices. Since $\leq$ relation is the partial order, the $\angle$ relation is a partial order. Therefore, the concept hierarchy generated using the superconcept and subconcept relations given in Definition 8 satisfies the requirements of a concept hierarchy given in Definition 7.

Figure 8 illustrates the hierarchical relations constructed from the conceptual clusters given in Figure 7. In Figure 7, the formal concept C_4, which is the supremum of the corresponding sublattice of the conceptual cluster CK_3, is a subconcept of C_1, which is a formal concept in the corresponding sublattice of the conceptual cluster CK_1. Therefore, CK_3 is a subconcept of CK_1. Similarly, hierarchical relations for other conceptual clusters are also generated to form the concept hierarchy that is shown in Figure 8. Each concept in the concept hierarchy is represented by a set of its attributes. The supremum and infimum of the lattice is considered as “Anything” and “Nothing” concepts, respectively.

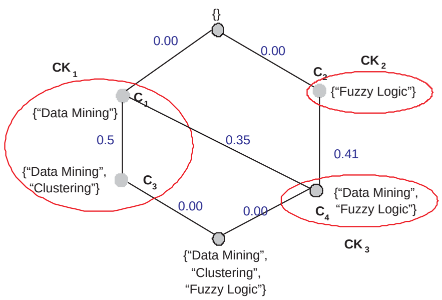

Fig. 7. Conceptual clusters.

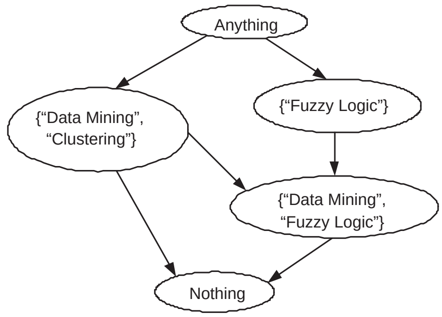

Fig. 8. Concept hierarchy.

7 Concept Hierarchy Generation from a Citation Database

We have applied the proposed approach to generate a concept hierarchy of research areas from a citation database. The citation database is created from a set of 1400 scientific documents on the research area "Information Retrieval" published in 1987-1997 downloaded from Institute for Scientific Information (ISI) [8]. The downloaded documents are preprocessed to extract related information such as the title, authors, journals and citation keywords before storing them into the citation database.

The construction of fuzzy formal context is performed as follows. For each document, we have extracted the 10 most frequent citation keywords. We then construct a fuzzy formal context $K_f = (G,M,I)$, with G as the set of documents and M as the set of keywords. The membership value of a document D on a citation keyword C_K in K_f is computed as

$$\mu(D, C_K) = \frac{n_1}{n_2}$$

where $n_1$is the number of documents that cited D and contained C_K and n_2 is the number of documents that cited D.

From the constructed fuzzy formal context, FFCA is used to generate the fuzzy concept lattice. Next, the Fuzzy Concept Clustering is applied to the fuzzy concept lattice to generate a set of conceptual clusters. Each cluster generated is considered as a research area in the main research theme "Information Retrieval". Each concept is represented by a set of keywords generated from documents belonging to the corresponding conceptual cluster. Then, Hierarchical Relation Generation is performed to construct a concept hierarchy of the research areas. The concept hierarchy reflects the taxonomy of research areas in the research theme "Information Retrieval", in which a research area can be a super-area or sub-area of other research areas.

The concept hierarchy can also be converted into ontology for sharing and reuse with other systems. Figure 9 depicts a part of the generated concept hier-

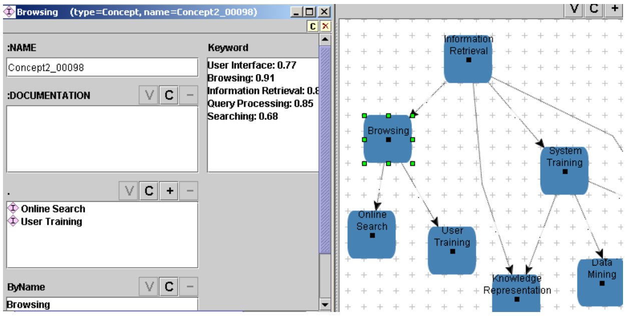

Fig. 9. An example of a concept hierarchy of research areas

archy of research areas with the similarity threshold $T = 0.7$. For simplification, we only use the keyword that has the highest membership value to label the research area. However, users can browse more detail information of each research area as illustrated in Figure 9.

8 Performance Evaluation

The performance of the concept hierarchy can be measured in order to evaluate the performance of the proposed fuzzy FCA-based approach for conceptual clustering. To do this, we use the *relaxation error* (RE) [2] to measure the goodness of the concepts generated. In addition, we also measure the *Average Uninterpolated Precision* (AUP) [12] to evaluate the retrieval performance from the concept hierarchy.

8.1 Evaluation Using Relaxation Error

To evaluate the goodness of the clusters generated, we measure the relaxation error, which implies dissimilarities of items in a cluster based on attributes' values. The relaxation error RE of a cluster C is defined as

$$RE(C) = \sum_{a \in A} \sum_{i=1}^{n} \sum_{j=1}^{n} P(x_i) P(x_j) d^a(x_i, x_j)$$

where A is the set of attributes of items in C, $P(x_i)$ is the probability of item x_i occurring in C and $d^a(x_i,x_j)$ is the distance of x_i and $x_j$on attribute a. In our application, $d^a(x_i,x_j) = |m(i,a) - m(j,a)|$ where $m(i,a)$ and $m(j,a)$ are the membership values of objects x_i and x_j on attribute a respectively. The cluster goodness G of cluster C is defined as

$$G(C) = 1 - RE(C)$$

As discussed in Section 7, we extract citation keywords of documents as their attributes. Since these attributes are used as descriptors for the generated clusters, we vary the number of keywords extracted to observe the effect of the keywords on cluster goodness. Besides, since COBWEB is considered as one of the most popular techniques for conceptual clustering, we also apply COBWEB to the citation database for performance comparison purposes. To use COBWEB, the membership values of keywords are replaced by appropriate nominal values. If the membership value is greater than 0.5, it is set as “Yes”, otherwise it is set as “No”.

Figure 10 shows the performance evaluation results on cluster goodness using FFCA and COBWEB while the number of extracted keywords is varied from 2 to 10. As shown in Figure 10, FFCA has achieved better cluster goodness than COBWEB. In addition, the experimental results have also shown that good cluster goodness is obtained when the number of extracted keywords is small. It is expected because smaller number of keywords used will cause smaller differences in objects in terms of keywords’ membership values. Therefore, the relaxation error will be smaller. However, as we will see later in Section 8.2, smaller number of extracted keywords will cause poor retrieval performance.

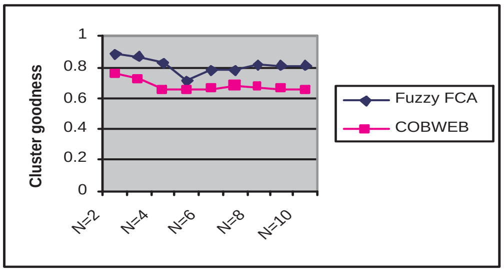

Fig. 10. Performance evaluation results on cluster goodness

8.2 Evaluation Using Average Uninterpolated Precision

The Average Uninterpolated Precision (AUP) is defined as the sum of the precision value at each point (or node) in a hierarchical structure where a relevant item appears, divided by the total number of relevant items. For evaluating AUP, we have manually classified the downloaded documents into classes based on their research themes. For each class, we extract 5 most frequent keywords from the documents in the class. Then, we use these keywords as inputs to form retrieval queries and evaluate the retrieval performance using AUP. This is carried out as follows. For each document, we will generate a set of *document keywords*. There are two ways to generate document keywords. The first way is to use the set of keywords, known as *attribute keywords*, from each conceptual cluster as the document keywords. The second way is to use the keywords from each document as the document keywords. Then, we vectorize the document keywords and the input query, and calculate the vectors' distance for measuring the retrieval performance.

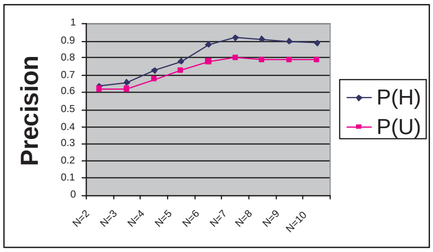

Fig. 11. Performance evaluation on precision.

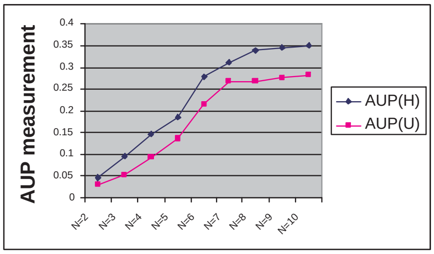

Fig. 12. Performance evaluation on AUP.

We refer the precision and AUP measured using the first way (i.e. using attribute keywords) to as *Hierarchical Precision* ($P(H)$) and *Hierarchical Average Uninterpolated Precision* ($AUP(H)$), as each concept inherits attribute keywords from its superconcepts. Whereas the precision and AUP measured using the second way (i.e. using keywords from documents) is referred to as *Unconnected Precision* ($P(U)$) and *Unconnected Average Uninterpolated Precision* ($AUP(U)$).

Figures 11 and 12 give the performance results for $P(H)$and $P(U)$and $AUP(H)$ and $AUP(U)$using different numbers of extracted keywords N. From Figure 11, we found that when N gets larger, the performance on $P(H)$, $P(U)$, $AUP(H)$and $AUP(U)$ gets better. When N is larger than 5, the values of $P(H)$ and $P(U)$ are considered as good performance (around 0.9 and 0.8 respectively). It has shown that the number of keywords extracted for conceptual clustering has affected the retrieval performance. In addition, the performance results on $P(H)$ and $AUP(H)$ are generally better than that of $P(U)$ and $AUP(U)$ respectively. It implies that the attribute keywords generated for conceptual clusters are more appropriate concepts for representing the concept hierarchical structure.

9 Conclusions

In this paper, we have proposed a fuzzy FCA-based approach for conceptual clustering for automatic generation of concept hierarchy from uncertainty information. The proposed approach consists of the following steps: Fuzzy Formal Concept Analysis, Fuzzy Conceptual Clustering and Hierarchical Relation Generation. In addition, we have also discussed an application that applies the proposed approach to generate a concept hierarchy of research areas from an experimental citation database. The performance evaluation of the proposed approach has also been presented based on the evaluation of the concept hierarchy generated from the citation database.

References

1. P. Cheeseman, J. Kelly, M. Self, J. Stutz, W. Taylor, and D. Freeman. AutoClass: A Bayesian Classification System. *Proceedings of the Fifth International Workshop on Machine Learning* Morgan Kauffmann, San Mateo, CA. 1988, pp. 54-64.
2. W. Chu and K. Chiang. Abstraction of High Level Concepts from Numerical Values in Databases. *Proceedings of AAAI Workshop on Knowledge Discovery in Databases*. 1994, pp. 133-144.
3. D.H. Fisher. Knowledge acquisition via incremental conceptual clustering. *Machine Learning. Vol. 2*. 1987, pp. 139-172.
4. B.Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundations*. Springer, Berlin - Heidelberg, 1999.
5. J.H Gennari, P. Langley and Fisher. *Models of Incremental Concept Formation, In J.Carbonell, Editor, Machine Learning: Paradigms and Methods.* The Netherlands: MIT Press,1990, pp. 11-62.
6. T.B. Ho. Incremental Conceptual Clustering in the Framework of Galois Lattice". In H. Lu, H. Motoda, H. Luu, editors.*KDD: Techniques and Applications, World Scientific*. 1997, pp. 49-64.
7. V.N. Huynh and Y. Nakamori. Fuzzy concept formation based on context model", in: N. Baba et al., editors. *Knowledge-Based Intelligent Information Engineering Systems & Allied Technologies*. IOS Press, Amsterdam, 2001, pp. 687-691.
8. ISI.Institute for Scientific Information. Available at: http://www.isinet.com, 2000.
9. I. Jonyer, D.J. Cook and L.B. Holder. Graph-based hierarchical conceptual clustering.*The Journal of Machine Learning Research, Vol. 2*. 2002, pp. 19-43.
10. C. Li and G. Biswas. Conceptual Clustering with Numeric and Nominal Mixed Data - A New Similarity Based System. *EEE Transactions on Knowledge and Data Engineering*. 1996.
11. Y. Lu. Concept Hierarchy in Data Mining: Specification, Generation and Implementation. *Technical Report*. Available at: http://gunther.smeal.psu.edu/3024.html.
12. N.Nanas, V.Uren and A. de Roeck. Building and Applying a Concept Hierarchy Representation of a User Profile. *Proceedings of the 26th annual international ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM Press, 2003.
13. S. Pollandt. *Fuzzy-Begriffe: Formale Begriffsanalyse unscharfer Daten, Springer Verlag*. Berlin - Heidelberg, 1996.
14. Y. Reich and S.J. Fenves. The formation and use of abstract concepts in design, In D.H. Fisher and M.J. Pazzani, Editors. *Concept Formation: Knowledge and Experience in Unsupervised Learning, Morgan Kaufmann*. 1991, pp. 323-353.
15. F. Vogt and R. Wille. TOSCANA: a Graphical Tool for Analyzing and Exploring Data, In R.Tamassia and I. G. Tollis, editors. *GraphDrawing' 94*. Heidelberg, 1995, pp.226-233.
16. L.A Zadeh. Fuzzy Sets. *Journal of Information and Control, Vol. 8*. 1965, pp. 338-353.