ISSN:0975-9646

G. Umamaheswara Rao et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (4) , 2012,4900-4903

Similarity of Fuzzy Ontology Generation for Semantic web

G. Umamaheswara Rao1,K.V. Subba Raju[2] ,K. Srikanth[3]

[1]M.V.G.R College of Engineering, M.Tech(CSE), Vizinagaram, Andhra Pradesh, India
[2]M.V.G.R College of Engineering, Assistance Professor, CSE Dept, Vizinagaram, Andhra Pradesh, India
[3]Andhra University, M.Tech.(P.hD), Computer Science and Systems Engineering, Visakapatnam, India

Abstract- First, Similarity Of Fuzzy Ontology Generation For Semantic Web, Fuzzy Formal Concept Analysis incorporates fuzzy logic into Formal Concept Analysis (FCA) to form a fuzzy concept lattice. The fuzzy ontology is an extension of the domain ontology for solving the uncertainty problems. Fuzzy Conceptual Clustering then constructs the concept hierarchy from the fuzzy concept lattice. Finally, Fuzzy Ontology Generation generates the fuzzy ontology from the concept hierarchy. We also discuss approximating reasoning for incremental enrichment of the ontology with new upcoming data. Finally, a fuzzy-based technique for integrating other attributes of database to the ontology is proposed. In this paper proposes a series of fuzzy ontology models that consist of fuzzy domain ontology and fuzzy linguistic variable ontologies, considering semantic relationships of concepts, including set relation, order relation and equivalence relation. Application of the fuzzy ontology to transportation knowledge modelling shows that this Similarity facilitates the knowledge share and reuse for fuzzy systems on the semantic web.

Keywords : Ontology Generation, Fuzzy Ontology. Formal Concept Analysis, Fuzzy Logic, Conceptual Clustering.

I. INTRODUCTION:

Ontology is a conceptualization of a domain into a human understandable, but machine-readable format consisting of entities, attributes, relationships and axioms [1]. Ontology uses classes to represent concepts. Ontology also supports taxonomy and non-taxonomy relations between classes. However, the conceptual formalism supported by typical ontology may not be scientific to represent uncertainty information that is commonly found in many application domains. For example, keywords extracted from scientific publications 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. In addition, it is sometimes difficult to judge whether a document belongs completely to a research area or not. To tackle this type of problems, one possible solution is to incorporate fuzzy logic into ontology to handle uncertainty data. Traditionally, fuzzy ontology is generated and used in text retrieval [2], in which membership values are used to evaluate the similarities between concepts on a concept hierarchy.

In this paper, we propose a framework known as FOGA (Fuzzy Ontology Generation framework) that can automatically generate a fuzzy ontology on uncertainty data. As compared with existing fuzzy ontology generation techniques, FOGA can automatically construct a hierarchy structure of ontology classes. In addition, this paper also discusses the use of FOGA to generate scholarly ontology for the Scholarly Semantic Web from an experimental citation database. Here, the taxonomy relations on ontology classes can be generated automatically as compared with the manual method used in other semantic scholarly systems such as ESKIMO [4]. However, FOGA still requires some minimal human interpretations to help add meaningful labels on initial class names, attributes and its relations.

II. RELATED WORK

Ontology Generation

Although editing tools [8], [9] have been developed to help users to create and edit ontology, it is a troublesome task to manually derive ontology from data. Typically, ontology can be generated from various data types such as textual data [10]. Compared to other types of data, ontology generation from textual data has attracted the most attention. Among techniques used for processing textual data, clustering is one of the most effective techniques for ontology learning.

III. FUZZY THEORY

In this section, we review some fundamental knowledge of fuzzy theory [3].

Definition 1 (Fuzzy Set). A fuzzy set A on a domain U, is defined by a membership function μ from U to [0.1], i.e., each item in A has a membership value given by μ We denote . φ(s) as a fuzzy set generated from a traditional set of items S. Each item in S has a membership value in [0, 1]. S can also be called as a crisp set.

Definition 2 (Fuzzy Relation). A fuzzy set A on a domain G ×M, where G and M are two crisp sets is a fuzzy relation on G, M.

Definition 3 (Fuzzy Sets Intersection). The intersection of fuzzy sets A and B, denoted as $A \cap B$, is defined by

$\mu_{A\cap B}(x)=\min(\mu_A(x), \mu_B(x))$.

Definition 4 (Fuzzy Sets Union). The intersection of fuzzy sets A and B, denoted as A U B, is defined by

$\mu_{AUB}(x) = \max(\mu_A(x), \mu_B(x))$.

Definition 5 (Fuzzy Set Cardinality). Let Sf be a fuzzy set on the domain U. The cardinality of Sf is defined as Where μ(x) is the membership of x in Sf .

$$|Sf| = \sum_{X \varepsilon \upsilon} \mu(x)$$

Definition 6 (Fuzzy Sets Similarity). The similarity between two fuzzy sets A and B is defined as

$$E(A,B) = \left| \frac{|A \cap B|}{AUB} \right|$$

Definition 7 (Fuzzy Sets Subsethood). The subsethood of a fuzzy set A of a conceptual cluster B is calculated as Subset

$$(A,B) = \left| \frac{|A \cap B|}{|B|} \right|$$

Definition 8 (Fuzzy Set Max-min Composition). Let P(X,Y) be a fuzzy relation on X ,Y and P (Y,Z) be a fuzzy relation on Y , Z. The max-min composition of P(X,Y) and Q(Y,Z), P • Q, is defined by:

$\mu_{P \cdot Q}(X,Y) = \max \min(\mu p(X,Y), \mu q(Y,Z))$,

$\varphi_x \ \varepsilon\ X, y\ \varepsilon\ Y$.

The max-min composition indicates the strength of relation between the element of X and Z.

IV. THE FOGA FRAMEWORK

Fig. 1 shows the proposed FOGA (Fuzzy Ontology Generation framework), which consists of the following components.

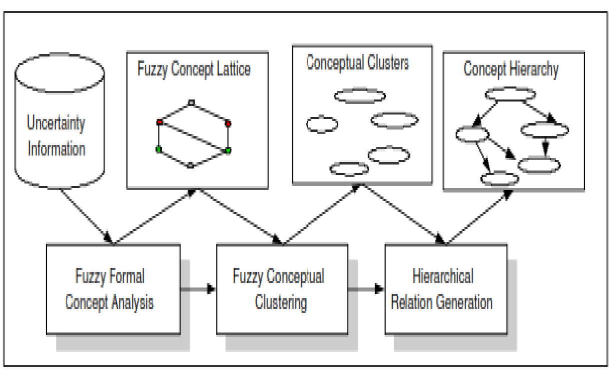

Fig 1. The FOGA framework.

A. Fuzzy Formal Concept Analysis

The Fuzzy Formal Concept Analysis incorporates fuzzy logic into Formal Concept Analysis to represent vague Information.

Figure 2 gives the traditional concept lattice generated from Table 1(a). Figure 3 gives the fuzzy concept lattice generated from the fuzzy formal context given in Table 1(b). As shown from the figures Fig.4.Ontology Generation Frame Work, 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.

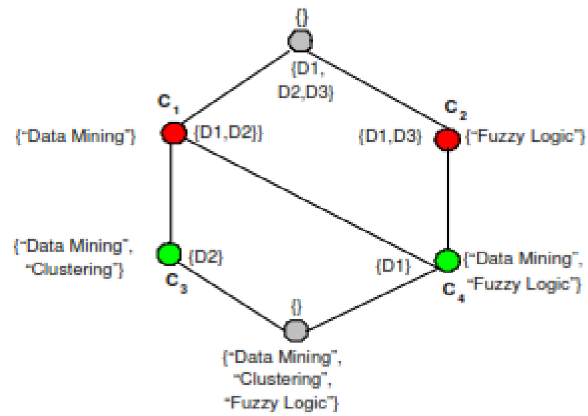

Fig. 2. A concept lattice generated from traditional FCA.

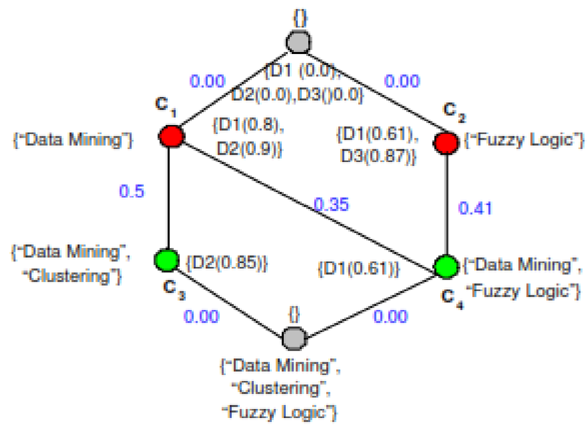

Fig. 3. A fuzzy concept lattice generated from FFCA.

Figure 2 gives the traditional concept lattice generated from Table 1(a). Figure 3 gives the fuzzy concept lattice generated from the fuzzy formal context given in Table 1(b). As shown from the figures Fig.4.Ontology Generation Frame Work, 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.

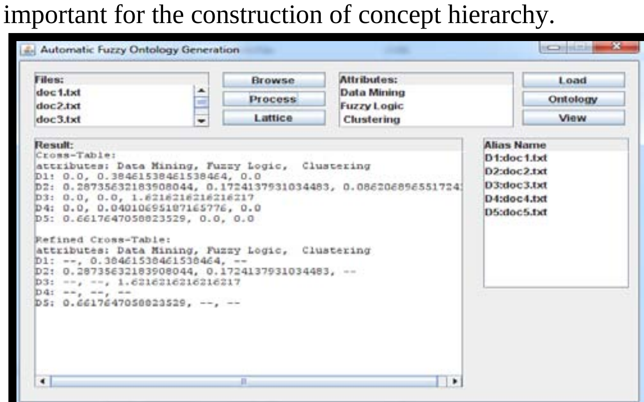

Fig.4.Ontology Generation Frame Work

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

TABLE1(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

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, "Data Mining" (D), "Clustering" (C) and "Fuzzy Logic" (F) representing three research topics. The relationship between an object and an attribute is represented by a membership value between 0 and 1.

V. Fuzzy Conceptual Clustering

Conceptual clusters have hierarchical relationships that can be derived from fuzzy formal concepts on the fuzzy concept lattice. That is, a concept represented by a conceptual cluster can be a sub concept or super concept of other concepts represented by other conceptual clusters. 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.

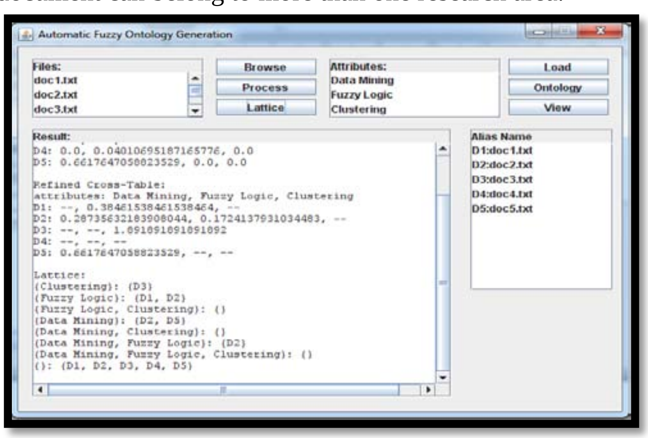

Fig. 5. Conceptual clusters.

Figure 5 shows the conceptual clusters that are generated from the concept lattice given in Figure 4 with the similarity confidence threshold T = 0.5. Figure 6 shows the corresponding concept Semantic Web, in which each concept is represented by a set of attributes of objects from the corresponding conceptual cluster.

VI. Semantic web

ontology generated for the Semantic Web contains information as a hierarchy of research areas as well as research areas for each document. Figure as a part of the generated Semantic Web of research areas. We use the keyword that has the highest membership value to label the research area. Nevertheless, users can browse more detail information of each research area.

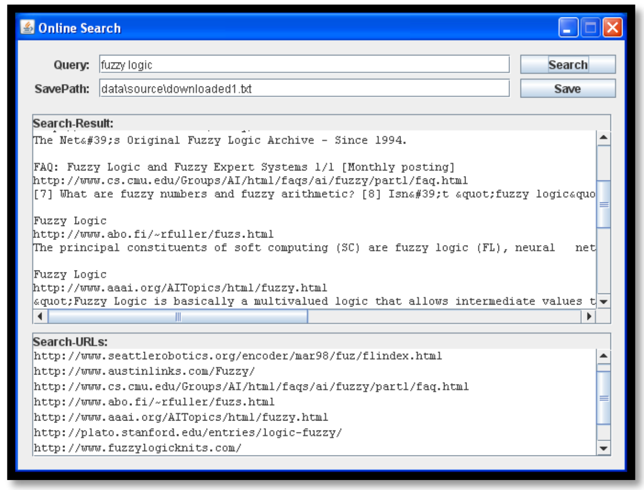

Fig.6. Semantic Web

TABLE 2

A Fuzzy Formal Context Having Cross Relation with the Fuzzy Formal Context in Table 2

	D1	D2	D3
Author1	1.0	-	1.0
Author2	0.5	1.0	-

The cross relation represents an intercontext relation that probably occurs between the fuzzy formal contexts when the set of objects of a context is regarded as the set of attributes of an other contexts. For example, the context represented by the cross table shown in Table 2 has cross relation with the context in Table 1(a) & 1(b), while the documents are used as attributes of the authors. The membership value of 1.0 implies that the author is the first author of the document, while 0.5 implies that the author is the second author.

VII. Performance evaluation

Generating Ontology from Citation Database To evaluate the proposed FOGA framework for ontology generation, we have collected a set of 1,400 scientific documents on the research area “Information Retrieval” published in 1987-1997 from the Institute for Scientific Information’s (ISI) Web site [52]. The downloaded documents are pre-processed to extract related information such as the title, authors, citation keywords, and other citation information. The extracted information is then stored as a citation database.

First, we construct a fuzzy formal context Kf ={G,M,I}, with G as the set of documents and M as the set of citation keywords. The membership value of a document D on a citation keyword CK in Kf is computed as $\mu(d,C_k) = n1/n2$, where n1 is the number of documents that cite D and contain CK and n2 is the number of documents that cite D. This formula is based on the premise that the more frequent akeyword occurs in the citing paper, the more important the keyword is in the cited paper.

A. Evaluation Using Recall, Precision and F-Measure

We have classified manually the documents downloaded from ISI into classes based on their research themes. These classes are used as a benchmark to evaluate the clustering results in terms of recall, precision, and F-measure. As discussed earlier, we extract citation keywords of documents as their attributes. Since these attributes are descriptors for the generated clusters, if more keywords are extracted and used, the more meaningful the cluster descriptors are constructed. To verify this, we vary the number of keywords N extracted from documents from 2 to 10, and the similarity threshold Ts from 0.2 to 0.9 when performing conceptual clustering. The measured precision, recall and F-measure are presented in Table 7, respectively.

TABLE 3:
Performance Results Using F-Measure Measurement

	T_s=0.2	T_s=0.3	T_s=0.4	T_s=0.5	T_s=0.6	T_s=0.7	T_s=0.8	T_s=0.9
N=2	0.78	0.78	0.78	0.78	0.77	0.76	0.76	0.76
N=3	0.79	0.79	0.79	0.79	0.77	0.76	0.76	0.76
N=4	0.83	0.86	0.86	0.87	0.82	0.79	0.78	0.78
N=5	0.84	0.85	0.85	0.86	0.83	0.81	0.81	0.81
N=6	0.85	0.85	0.86	0.86	0.84	0.82	0.82	0.82
N=7	0.88	0.86	0.87	0.87	0.86	0.85	0.85	0.85
N=8	0.86	0.86	0.86	0.87	0.85	0.85	0.84	0.84
N=9	0.84	0.84	0.86	0.86	0.85	0.84	0.83	0.83
N=10	0.81	0.82	0.84	0.84	0.83	0.83	0.83	0.83
Average	0.83	0.83	0.83	0.84	0.82	0.81	0.8	0.8

VIII. CONCLUSIONS

The generated ontology represents knowledge on documents and its research areas. The performance evaluation of the proposed FOGA framework has also been given based on the generation of the scholarly ontology. Fuzzy Formal Concept Analysis, Fuzzy Conceptual Clustering, Fuzzy Ontology Generation, and Semantic Web Representation Conversion. In addition, we have also proposed an approximating reasoning technique that allows the generated fuzzy ontology to be incrementally furnished with new instances. Finally, we have also proposed a technique to integrate extra attributes in a database to the ontology. Our authoring tool provides Similarity support for the customization of dynamic web documents based on comparing the pages generated by the system with a modified version provided by the end-user. DESK is based on PEGASUS, a system used to represent the Semantic Web information structured by models that allow a clear separation between contents and presentation. DESK uses domain information stored in PEGASUS and presentation models for finding the context of changes made by user. Our authoring tool also determines whether the user is enabled to do these modifications depending on a user model. With DESK the user only needs to take care of editing HTML pages using any standard HTML editing tool such as PageMaker or Netscape Composer.

REFERENCES

[1] N. Guarino and P. Giaretta, Ontologies and Knowledge Bases:Towards a Terminological Clarification. Toward Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing. Amsterdam: IOS Press, 1995.

[2] Castells, P., Szekely, P.: Presentation Models by Example. En: Duke, D.J., Puerta A. (eds.): Design, Specification and Verification of Interactive Systems '99. Springer-Verlag, 1999, pp. 100-116.

[3] L.A. Zadeh, “Fuzzy Logic and Approximate Reasoning,” Synthese, vol. 30, pp. 407-428, 1975.

[4] Castells, P. Macías, J.A.: An Adaptive Hypermedia Presentation Modeling System for Custom Knowledge Representations. Proceedings of the World Conference on the WWW and Internet

[5] D.H. Widyantoro and J. Yen, “A Fuzzy Ontology-Based Abstract Search Engine and Its User Studies,” Proc. 10th IEEE Int’l Conf. Fuzzy Systems, pp. 1291- 1294, 2001.

[6] B. Ganter and R. Wille, Formal Concept Analysis: Mathematical Foundations. Springer, 1999. [7] W3C, “Web Ontology Language Overview,”http://www.w3.org/ TR/owl-features/, 2006.

[7] OntoWeb develiverable 1.3, http://www.ontoweb.org/, 2006.

[8] A. Gomez-Perez, O. Corcho, and M. Fernandez-Lopez, Ontological Engineering: With Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web (Advanced Information and Knowledge Processing). Springer, 2004.

[9] Macías, J. A., Castells, P.: Adaptive Hypermedia Presentation Modeling for Domain Ontologies. To appear in proceedings of 10th International Conference on Human-Computer Interaction (HCII ’2001). New Orleans, Louisiana, 2001.

[10] Murray, T.: Authoring Kowledge Based Tutors: Tools for Content, Instructional Strategy, Student Model, and Interface Design. Journal of the Learning Sciences 7, 1, 1998, 5-64.

[11] Muslea, I.: Extraction Patterns for Information Extraction Tasks: A Survey. Proceedings of AAAI Workshop on Machine Learning for Information Extraction. Orlando, Florida, july 1999.

[12] Vassileva, J.: Dynamic Course Generation on the WWW. Actas 8th World Conference on Artificial Intelligence in Education (AIED’97). Kobe, Japón, 1997, 498-505.

[13] Weber, G. and Specht, M.: User modeling and Adaptive Navigation Support in WWW-based Tutoring Systems. Proceedings 6th International Conference on User Modeling (UM97). Sardinia, Italy, 1997.