Constructing and Merging Ontology of E-business based on Fuzzy Rough Concept Sets

Yan Wang, Hong-Sheng Xu

Advances in Information Sciences and Service Sciences, Volume3, Number7, August 2011

Abstract

Ontology, regarded as a basic knowledge structure for information on the Semantic Web, often captures a semantic network and represents it by a graph. The domain ontology model of e-business is built combined with knowledge of domain expert. The domain ontology model of e-business is built combined with knowledge of domain expert. Ontology provides common understanding of the domain knowledge and confirms common approbatory vocabulary in the domain. The project probes into building and merging of e-business domain ontology based on fuzzy rough concept sets in order to constructing the ontology model of e-commerce recommendation system, and to suffice the needs of theory and application in E-commerce recommendation system. Finally, ontology merging is carried out. The experimental results indicate that this method has great promise.

Keywords: Domain Ontology, Variable Precision Rough Set, Fuzzy Sets, Ontology Merging

1. Introduction

As the foundation of the semantic web, ontology is a formal, explicit specification of a shared conceptual model. The aim of ontology is to obtain, describe and express the knowledge of related domain. The domain ontology model of e-business is built combined with knowledge of domain expert. The project probes into building and merging of e-business domain ontology based on rough concept lattices in order to constructing the ontology model of e-commerce recommendation system, and to suffice the needs of theory and application in E-commerce recommendation system[1]. Rough concept lattice model is proposed based on integrating of variable precision rough set model, formal concept analysis and the theory of fuzzy sets, and is used to reduce formal context. The method of conceptual similarity in electronic commerce based on FRFCA is proposed by introducing of trapezoidal fuzzy number to represent the fuzzy similarity and via weighted composite, non-fuzzy, computing the composite similarity among multi-strategies heterogeneous ontology concepts for the sake of improving the effect of mapping and matching.

Ontology, regarded as a basic knowledge structure for information on the Semantic Web, often captures a semantic network and represents it by a graph whose nodes denote concepts or individual objects and arcs represent the relationships or associations among concepts. Ontologies allow web resources to be semantically enriched. This is a pre-condition to provide new, advanced services over the web, such as semantic search and retrieval of web resources. Ontology building is a task that pertains to ontology engineers, an emerging expert profile that requires the expertise of knowledge engineers (KEs) and domain experts (DEs). Such an ontology can play the role of a unifying theory of computer-based situation awareness. In this paper we describe all the concepts in this ontology. Ontology, which is a theory about the nature of existence in philosophy, plays an important role in the integrating heterogeneous information.

Concepts and relationships are basic components in an ontology. Web documents are the most important source for deriving concepts and relationships. Ontology provides common understanding of the domain knowledge and confirms common approbatory vocabulary in the domain, as well as gives specific definition of the relation between these vocabularies from formal model of different levels. Therefore, it becomes a pivotal issue on ontology application to build ontology. But this field is still being discussed today. According to the goal of ontology, the pivotal problem of ontology constructing is to find the concepts and relationship among concepts after confirming the field, but these are connotative in the brain or store the file of this field in actual life. An ontology is usually defined as the
specification of a conceptualization. Ontology is an emerging discipline to address semantic issues and enhance knowledge management applications[2]. With the assistance of ontological technologies, users can retrieve knowledge in a semantic manner. The core features of an ontology are a set of concepts, a set of properties and the relationships between the elements of these two sets. Ontology and metadata are widely accepted as the core elements of the semantic web. The ontology multiple-layered framework corresponds to the semantic layer of the traditional semantic web stack. The meta layer contains rdfs:Class only. The rdfs:Class is the root class in the RDF data model. Any other class is regarded as an instance of the rdfs:Class. RDF, RDFS. Ontology is field of philosophy that is the study of being and existence. In information technology, ontology has been proposed as an explicit specification for conceptualization.

Rough set theory is a way of representing and reasoning imprecision and uncertain information in data. We will take a critical look at the Pawlak and probabilistic rough set models in this section. Variable precision rough set model (VPRS), as introduced by Ziarko in 1993, is one of the most important branches. Standard inclusion relation is extended to majority inclusion relation in VPRS. Variable precision rough set (VPRS) model is studied in this paper, include β value select algorithm , algorithm of reduction attribute and extraction rule. The relation between β threshold value of reliability and approximate quality of classification is analyzed firstly, and one algorithm confirm the range of β value by the r threshold value of approximate quality of classification are provided. The algorithm integrates both the fuzzy set theory and the variable precision rough-set model to discover fuzzy knowledge. The fuzzy β-certain rules with misclassification degrees smaller than β and the fuzzy β-possible rules with misclassification degrees smaller than 1 − β are derived. Compared to traditional rough-set based induction algorithms, the proposed approach has an additional conversion procedure. The variable precision rough-set model can be thought of as a generalization of the rough-set model. It allows for some degree of uncertainty and misclassification in the mining process.

The paper probes into constructing and merging domain ontology of e-business based on fuzzy rough concept sets in order to constructing the ontology model of e-commerce recommendation system, and to suffice the needs of theory and application in E-commerce recommendation system. The paper offers a methodology for building ontology and carries on ontology merging for knowledge sharing and reusing based on concept lattice union. Finally, ontology merging is carried out.

2. Fuzzy rough concept sets mode

Dubois and Prade first introduced the concept of fuzzy rough sets and constructed a pair of fuzzy rough approximations of a fuzzy set by using the notions of the greatest t-norm (min), its dual t-conorm (max), and a fuzzy similarity relation (a fuzzy relation with reflexivity, symmetry and transitivity). The new fuzzy rough approximation operators are established based on a fuzzy covering, a binary fuzzy conjunction logical operator with lower semi-continuity in its second argument, and the adjunctional implication operator of the conjunction. Fuzzy set theory and fuzzy logic, proposed by Zadeh, are acknowledged as an appropriate formalism for capturing imprecise and vague knowledge. While in classical set theory elements either belong to a set or not, in fuzzy set theory elements can belong to a certain degree. Fuzzy rule-based systems have some advantages over other formalisms: they provide a natural representation for human knowledge as well as a very interpretable model (since the semantics of the rules can easily be understood even for not experts users), are simpler, cheaper and more robust than their crisp versions and, last but not least, have shown to behave very well in practical applications. Properties of existent fuzzy concept lattices derived from an adjoint pair of operations are first reviewed and examined. Based on both lattice-theoretic and fuzzy set-theoretic operators, two new pairs of fuzzy rough set approximations within fuzzy formal contexts are then defined.

This paper presents a comparative study of concept lattices of fuzzy contexts based on formal concept analysis and rough set theory[3]. It is known that every complete fuzzy lattice can be represented as the concept lattice of a fuzzy context based on formal concept analysis. The notions of partial order, lattice order, and formal concept are generalized for fuzzy setting. Presented is a theorem characterizing the hierarchical structure of formal fuzzy concepts arising in a given formal fuzzy context.
Let $K = (X, Y, I)$ be an $L$-context (fuzzy context). We define an associated underlying fuzzy graph $G = (X \cup Y, u, p)$, where $u$ is the membership function of $X \cup Y$ into $L$ and $p$ is the membership function of edges such that $u(x) \neq 0$, $u(y) \neq 0$ with $\text{supp}(u(x))$, $\text{supp}(u(y))$ complete graphs, $\forall x \in X$, $\forall y \in Y$.

Assume the sketch map for Trapezoidal Fuzzy is given by experts as shown in figure 1.

![Figure 1. Trapezoidal Fuzzy sketch map](image)

Let $A \subseteq EM$. $A$ is a set of fuzzy sets selected in advance to be possible features for data clustering problem on $X$. In the following, we try to extend the above algorithm for crisp attributes to the fuzzy situation within the framework of fuzzy concept lattices.

$$\text{LF} (I_i) = \frac{F_3(I_i)}{\sum_{j=1}^{n} F_3(I_i) \{F_3(I_1), F_3(I_2), \ldots, F_3(I_n)\}}$$

(1)

Step 1: Find fuzzy set. For each $x \in X$, every variable in the consequent can only be used in one family of rules.

Step 2: we use $\zeta_i$, the fuzzy description of each $x \in X$, to establish the fuzzy relation matrix $M=(m_{ij}), m_{ij} = \min\{\mu_{A_1} \ A_{C_1}(x_1), \mu_{A_2} \ A_{C_2}(x_2)\}$, which is $\delta x$ in the crisp situation, by the following computations.

The method of conceptual similarity in electronic commerce based on FRFCA is proposed by introducing of trapezoidal fuzzy number to represent the fuzzy similarity and via weighted composite, non-fuzzy, computing the composite similarity among multi-strategies heterogeneous ontology concepts for the sake of improving the effect of mapping and matching. The relation between $\beta$ threshold value of reliability and approximate quality of classification is analyzed firstly, and one algorithm confirm the range of $\beta$ value by the $r$ threshold value of approximate quality of classification are provided[4]. The intersection of two fuzzy sets $A$ and $B$ is denoted by $A \cap B$, and the membership function of $A \cap B$ is given by.

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

(2)

We propose an approach, not only to reducing fuzzy attributes, but also to inducing fuzzy rules. A technique that is not currently employed in video deinterlacing is the fuzzy-rough sets hybrid scheme, which has been researched by many authors.

Input: Every variable in the antecedent can only be used in one family of rules, As in the comparison with FuzzyK, we set the number of clusters to be 32 for both SOM and Gaussian mixture model, with the maximum number of iterations of 100,000 and 100.

Output: variables are defuzzified using the centroid of area, $R$ is the set of all $\beta$-reducts in $B$ with respect.

Stage 1: Partition the object set into disjoint subsets according to class labels. Denote each set of objects belonging to the same class $C_1$ as $X_1$.

Stage 2: If threshold $\alpha = 1$, $C_1 = \{x_1\}$, then fuzzy description $\zeta_{C_1} = \{m_6, m_9\}$.

Step 3: If such a set $N_i \subseteq P$, which has I elements, exists, go to Step 7; Otherwise, go to Step 6.
Step 4: Select a set \( M_i \subseteq P \) which has \( l \) elements. Let \( H' = H \cup M_i \).

Step 5: Intuitively, this concept measures if the co-occurrence of \( x \) and \( y \) \( (P(x, y)) \) is more likely than their independent occurrences \( (P(x) \cap P(y)) \).

Step 6: If \( P \neq \emptyset \), go to Step 4; Otherwise, Stop.

The values in the previous step can fire some rules. From the input variables, intermediate variables are calculated. Then, intermediate variables are used to compute another intermediate variables and, finally, output variables. The algorithm calculates the fuzzy lower and the fuzzy boundary approximations of single attributes from the terminal level to the higher level. After that, the fuzzy lower and the fuzzy boundary approximations of more than one attribute are derived based on the results of single attributes. This paper revisits attribute reduction based on fuzzy variable precision rough sets. Different types of fuzzy approximation operators can be defined by using two pairs of fuzzy approximation operators and various t-norms.

3. Ontology building and merging based on fuzzy rough concept sets

The paper probes into building and merging domain ontology of e-business based on fuzzy rough concept lattices in order to constructing the ontology model of e-commerce recommendation system, and to suffice the needs of theory and application in E-commerce recommendation system.

3.1. Constructing ontology method based on fuzzy rough concept lattice

According to the goal of ontology, the pivotal problem of ontology constructing is to find the concepts and relationship among concepts after confirming the field, but these are connotative in the brain or store the file of this field in actual life. The primary content of this dissertation is to apply fuzzy rough formal concept analysis technology to obtain all connotative concepts and hierarchy of them automatically from the designated data, which is not under the influence of developer. Denoting concept in symbol achieves formalized conceptual model. Before doing this, we introduce two new concepts to compare two types of attribute reduction. A concept similarity measure method based on fuzzy rough concept lattice is proposed to enhance quality of recommendation when lacking for sort of user’s interests. The experimental results show that this method can effectively improve the performance of the recommendation system. We describe and exemplify the basic concepts of knowledge representation in our proposed fuzzy rough approach, specifically the lower and upper approximations, fuzzy rough set, positive region, negative region and boundary region[5].

\[
\text{If } M_i \subseteq M, \text{ then we say that } M \text{ is } \beta\text{-definable in } \alpha\text{-approximation. Otherwise, we say that } M \text{ is } \beta\text{-undefinable in } \alpha\text{-approximation, and } M \text{ is a fuzzy-rough set.}
\]

The fuzzy rough set technique, which combines rough set theory and fuzzy set theory, is not only able to describe indiscernibility among objects but can also handle the fuzziness of objects.

\[
f_{jk} = \text{relevancy}(T_j, T_k) = \frac{\sum_d d_{ijk} \times \text{WeightingFactor}(T_k)}{\sum_d d_{ij}}
\]

Here the type of attribute reductions is the set of all the attribute reductions obtained by certain type of fuzzy approximation operator. A concept similarity measure method based on fuzzy rough concept lattice is proposed to enhance quality of recommendation when lacking for sort of user’s interests, i.e, definition 3.

IF \( p(I_1, I_2) = \{<a_1, b_2>, \ldots, <a_n, b_n>\} \), Then we say the type of attribute reductions \( \text{reduct}_{type_1} \) and the type of attribute reductions \( \text{reduct}_{type_2} \) are identical.

\[
d_{ijk} = \text{tf}_{ijk} \times \log_{10} \left( \frac{N}{df_{jk}} \times w_j \right) \quad d_{ij} = \text{tf}_{ij} \times \log_{10} \left( \frac{N}{df_j} \times w_j \right)
\]
Definition 4 shows that in the case of given $T$-transitivity of the fuzzy $T$-similarity relation of e-business, the triangular norms in the $S$-lower approximation operator have no effect on the result of attribute reduction based on fuzzy rough concept lattice. We find that three different types of lower approximation constructed by the similarity relation with the same type of transitivity have the identical reductions of fuzzy rough concept lattice. But their work was restricted to the fuzzy $T$-rough sets defined by the fuzzy $T$-similarity relation and their lower and upper approximation operators are not dual.

### 3.2. Ontology merging based on fuzzy rough concept sets

The output formats vary from fully merged ontologies to class-to-class mappings[6]. The tools also provide different levels of support to users and require varying amounts of human input. Many of the tools merge ontologies without human intervention; others (like PROMPT) suggest matches and require an expert to verify them at each step in the merging/mapping process. This is usually considered better for most mapping scenarios.

The paper offers a methodology for building ontology and carries on ontology merging for knowledge sharing and reusing based on fuzzy rough concept sets union. Moreover, ontology is presented by expressing similar relation of concepts in non-hierarchy and the relevancy of concepts and objects in combination with probabilistic model. It is a common practice that, during the process of concept mapping, we need to evaluate the similarities of two concepts for different text features as a set of intermediate solutions. Therefore, the effect of the text information may be excessively emphasized, and the structural information contribution to the similarity computation may be weakened. In order to improve the mapping effect, fuzzy similarity method is imported to denote different similarities based on diversified features.

It was defined as the amount of information provided by the occurrence of an event (y) of another event (x) and formulated as $A = (a_1, \cdots, a_m)$, $a_i = (DataAttributeSet_i, ObjectAttributeSet_i)$, $a_j = (DAIns\ tan\ ceSet_{i1}, \cdots, DAIns\ tan\ ceSet_{im})$. Given the threshold $\beta$, $\beta \in [0, 1]$, the set $P$ consists of the fuzzy attributes whose indispensability degrees in $B$ are greater than $1 - \beta$, $(OAReferenceConcept_{i1}, \cdots, OAReferenceConcept_{in})$ are the set of all fuzzy condition attribute-values and the set of all fuzzy decision attribute-values of the rule $X_i$, respectively. $H(x_i)$ is the set of fuzzy attribute-values which are $\beta$-indispensable in the rule $x_i$. For every attribute $m \in F$, $m$ is a crisp or fuzzy attribute which is objectively determined by the given data and facts. We study the classification with $A=\{male, engineer, lawyer\} \subseteq F$ based on the data shown in table 1.

<table>
<thead>
<tr>
<th>II</th>
<th>humidity</th>
<th>caves</th>
<th>volcano</th>
<th>Woods</th>
<th>tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doc1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Doc2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Doc3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Doc4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Doc5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
In addition, due to the changing environment of dynamic applications, the feature contents of an ontology concept may also change occasionally. The fuzzy decision table requires the reduction, not just of fuzzy attributes, but also of the superfluous fuzzy attribute-values. Reducing fuzzy attribute-values is equivalent to reducing every initial fuzzy rule.

This approach classifies pixels into several categories that are differentiated according to the value of the fuzzy logic principles. The statistics of the ontologies developed for individuals and vehicles from the three data sources are shown in table 2.

<table>
<thead>
<tr>
<th>Doc4</th>
<th>forest</th>
<th>wind breaks</th>
<th>timber</th>
<th>forest farm</th>
<th>tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doc5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Doc6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Doc7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Doc8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. The formal context of environment forest

From the monotone, fuzzy similarity or extensivity of the generalized fuzzy rough approximation operators, the inequalities $T = (t_1, \cdots, t_n)$ and $t = (\text{ParentConcept}, \text{ChildrenConceptSet})$ hold. The following proposition shows further results for hybrid compositions of the generalized fuzzy rough approximation operators. In contrast with the complement, inclusion, intersection and union of classical fuzzy sets, the corresponding operations of lattice fuzzy sets should be performed with a negation, a partial order, infimum and supremum, respectively.

The paper offers a methodology for carrying on ontology merging for knowledge sharing and reusing based on fuzzy rough concept lattice union. Moreover, ontology is presented by expressing similar relation of concepts in non-hierarchy and the relevancy of concepts and objects in combination with probabilistic model.

A special notation is often used in the literature to represent fuzzy sets. Assume that $x_1, \cdots, x_n$ are the elements in fuzzy set $A$, and $\mu_1, \cdots, \mu_n$ are, respectively, their grades of membership in $A$. We, therefore, generalize them to fuzzy relative reduct and fuzzy relative core, which is also the generalization of the traditional relative reduct and relative core. In this section, we describe and exemplify basic concepts of fuzzy relative attribute reduction that is the inconsistence degree between two samples, the relative indispensability degree, fuzzy relative reduct and fuzzy relative core.

4. Experimental results

The paper probes into building and merging domain ontology of e-business based on fuzzy rough concept lattices in order to constructing the ontology model of e-commerce recommendation system, and to suffice the needs of theory and application in E-commerce recommendation system. To demonstrate the effectiveness of the proposed algorithm, we used it to classify Data containing 150 training instances. The proposed algorithm was run on a training set to induce fuzzy $\beta$-certain and $\beta$-possible rules. The rules derived were then tested on the test set to measure the percentage of correct predictions. Classification rates were then averaged across all possible groups. The proposed methods were implemented on a Pentium IV processor (3.60 GHz). Experiments were conducted to evaluate the performance of the proposed FRFCA method. To demonstrate the effectiveness of the proposed algorithm, we used it to classify Fisher’s Iris Data containing 150 training instances. We use some numerical datasets from UCI machine learning repository. Figure 3 shows that the method of isomorphic generating of ontology merging based on rough concept lattices is presented.
In the following, we analyze the comparison results in three areas: the capability to find both reduct and the quality of rule concept, i.e. classification performance and size of fuzzy rule lattices, and the time complexity. The ontology context can be split onto ontology block in smaller order of property, and searching a isomorphic ontology in the ontology context kernel with the same order.

5. Conclusions

The paper probes into building and merging domain ontology of e-business based on fuzzy rough concept lattices in order to constructing the ontology model of e-commerce recommendation system, and to suffice the needs of theory and application in E-commerce recommendation system. Finally, ontology merging is carried out. In applications, we have studied a fuzzy algorithm and a type of fuzzy descriptions of objects in which the attributes of objects are described for databases with fuzzy rough concept lattices. The experimental results for the given illustrative examples demonstrate the effectiveness of the theoretical contributions proposed in this paper.

6. References