

Fuzzy Ontology Models Based on Fuzzy Linguistic Variable for Knowledge Management and Information Retrieval

Jun Zhai, Yiduo Liang, Jiatao Jiang and Yi Yu

School of Economics and Management, Dalian Maritime University, Dalian, 116026, China

Abstract: Ontology is the basis of sharing and reusing knowledge on the Semantic Web, and ontology-based semantic retrieval is a hotspot of current research. Fuzzy ontology is an extension of domain ontology for solving the uncertainty problems. To represent fuzzy knowledge more effectively, this paper presents a new series of fuzzy ontology models that consists of fuzzy domain ontology and fuzzy linguistic variable ontologies, considering semantic relationships of concepts, including set relation, order relation, equivalence relation and semantic association relation etc. The process to construct linguistic variables ontology is discussed. Using ontology and RDFS, the knowledge model for product information is created. To achieve semantic retrieval, the semantic query expansion in SeRQL is constructed by semantic relations between fuzzy concepts. The application shows that these models can overcome the localization of other fuzzy ontology models, and this research facilitates the fuzzy knowledge sharing and semantic retrieval on the Semantic Web.

1. Introduction

Ontology is a conceptualization of a domain into a human understandable, machine-readable format consisting of entities, attributes, relationships, and axioms [1]. It is used as a standard knowledge representation for the Semantic Web. However, the conceptual formalism supported by typical ontology may not be sufficient to represent uncertainty information commonly found in many application domains due to the lack of clear-cut boundaries between concepts of the domains. Moreover, fuzzy knowledge plays an important role in many domains that face a huge amount of imprecise and vague knowledge and information, such as text mining, multimedia information system, medical informatics, machine learning, and human natural language processing.

To handle uncertainty of information and knowledge, one possible solution is to incorporate fuzzy theory into ontology. Then we can generate fuzzy ontologies, which contain fuzzy concepts and fuzzy memberships. The fuzzy ontologies are capable of dealing with fuzzy knowledge [2], and are efficient in text and multimedia object representation and retrieval.

Lee et al. proposed an algorithm to create fuzzy ontology and applied it to news summarization [3]. This work is based on their previous work on ontology-based fuzzy event extraction agents for Chinese news summarization. Tho et al. proposed a Fuzzy Ontology Generation Framework (FOGA) for fuzzy ontology generation on uncertainty information [4]. This framework is based on the idea of fuzzy theory and Formal Concept Analysis (FCA). Abulaish et al. proposed a fuzzy ontology framework in which a concept descriptor is represented as a fuzzy relation which encodes the degree of a property value using a fuzzy membership function [5]. To enable representation and reasoning for fuzzy ontologies, Kang et al. proposed a new fuzzy extension of description logics called the fuzzy description logics with comparison expressions (FCDLs) [6]. Calegari et al. presented the fuzzy OWL language [7]. Zhai et al. studied the fuzzy ontology using intuitionistic fuzzy set [8].

But, current fuzzy ontology models have localization in fuzzy semantic retrieval. To represent formally the fuzzy knowledge more effectively, this paper presents a new kind of fuzzy ontology models. The rest of this paper is organized as follows: Section 2 introduces fuzzy domain ontology model. Section 3 presents fuzzy linguistic variable ontology models. Section 4 applies fuzzy ontology to knowledge modeling and section 5 studies information retrieval based on SeRQL. Finally, section 6 concludes the paper.

2. Fuzzy domain ontology model

An ontology organizes domain knowledge in terms of concepts, properties, relations and axioms, and the fuzzy ontology is created as an extension to the standard ontology.

Definition 1 (Fuzzy domain ontology) – A fuzzy domain ontology is a 6-tuple $O_F = (I, C, P^C, R, P^R, A_F)$, where:

- (1) I is the set of individuals, also called instances of the concepts.
- (2) C is a set of concepts. A concept is often considered as a class in an ontology. Every concept here has some properties whose value is fuzzy concept or fuzzy set. And, every concept can have the degree of membership $\mu_C(i) : I \rightarrow [0,1]$ and the degree of non-membership $\nu_C(i) : I \rightarrow [0,1]$ of the $i \in I$ in C .

(3) P^C is a set of concepts properties. A property $p^C \in P^C$ is defined as a 5-tuple of the form $p^C(c, v_F, q_F, f, U)$, where $c \in C$ is an ontology concept, v_F represents property values, q_F models linguistic qualifiers, which can control or alter the strength of a property value v_F , f is the restriction facets on v_F , and U is the universe of discourse. Both v_F and q_F are the fuzzy concepts on U , but q_F changes the fuzzy degree of v_F . For example, “price” is a property of concept “product”. The value of “price” may be either fuzzy concept “cheap” or fuzzy number “around 50”, and the linguistic qualifiers may be “very”, “little”, “close to” etc. Therefore, the final value of “price” may be “very cheap” or “little expensive”. At the same time, the property $p^C \in P^C$ has also the non-fuzzy form $p^C(c, v, f)$.

(4) R is a set of inter-concept relations between concepts. The relation type is not only the ordinary binary relation of $r \subseteq C \times C$, but also is the fuzzy relation and the intuitionistic fuzzy relation from C to C .

(5) P^R is a set of relations properties. Like concept properties, $p^R \in P^R$ is defined as a 4-tuple of the form $p^R(c_1, c_2, r, s_F)$, where $c_1, c_2 \in C$ are ontology concepts, r represents relation, and $s_F \in [0,1]$ or $s_F \subseteq [0,1]$ models relation strengths and has meaning of fuzzy set or intuitionistic fuzzy set on $C \times C$, which can represent the strength of association between concept-pairs $\langle c_1, c_2 \rangle$. For instance, there is a relation of “loyalty” between “customer” and “brand”. The strength of “loyalty” can be 0.7, a fuzzy value, and can be [0.6,0.8], a interval value, i.e. intuitionistic fuzzy value, which express more abundant information about uncertainty. On the other hand, s_F can also be fuzzy linguistic value, i.e. fuzzy concept.

(6) A_F is a set of fuzzy rules. In a fuzzy system the set of fuzzy rules is used as knowledge base.

The fuzzy domain ontology is used to model domain expert knowledge. But, due to the lack of relationships between fuzzy concepts that can be the value of properties, it is difficult to integrate diverse ontology systems. For example, in an ontology the set of property “price” value is {cheap, appropriate, expensive, ...}, and in other ontology the same set is {high, low, middle, ...}. To map these ontologies, it is necessary to define the semantic relationship between fuzzy concepts, e.g. “cheap” and “expensive” have the relation of disjointness, and “low” and “high” have the same relation of disjointness etc.

Consequently, we propose the fuzzy linguistic variables ontology models.

3. Fuzzy linguistic variable ontology

The fuzzy linguistic variables proposed by Zadeh are the basic of fuzzy knowledge and fuzzy system. To achieve the knowledge share and reuse for fuzzy systems on the Semantic Web, it is necessary to represent the fuzzy linguistic variables with ontology.

Definition 2 (Fuzzy linguistic variable) – Fuzzy linguistic variable is the variable whose value is term or concept in natural language. A fuzzy linguistic variable is a 4-tuple (X, T, M, U) , where:

- (1) X is the name of fuzzy linguistic variable, e.g. “price” or “speed” etc.
- (2) T is the set of terms which is the value of fuzzy linguistic variable, e.g. $T = \{\text{cheap, appropriate, expensive, ...}\}$ or $T = \{\text{fast, middle, slow, ...}\}$.
- (3) M is the mapping rules which map every term of T to fuzzy set at U .
- (4) U is the universe of discourse.

Introducing semantic relationships between concepts, we obtain the ontology model.

Definition 3 (Fuzzy linguistic variable ontology) – A fuzzy linguistic variable ontology is a 6-tuple $O_F = (c_a, C_F, R, F, S, U)$, where:

(1) c_a is a concept on the abstract level, e.g. “price”, “speed” etc. The corresponding element of c_a is X in definition 2.

(2) C_F is the set of fuzzy concepts which describes all values of c_a . The corresponding element of C_F is T in definition 2, but C_F has certain structure or relations R .

(3) $R = \{r \mid r \subseteq C_F \times C_F\}$ is a set of binary relations between concepts in C_F . A kind of relation is set relation $R_S = \{\text{inclusion (i.e. } \subseteq \text{), intersection, disjointness, complement (i.e. } \neg \text{)}\}$, and the other relations are the order relation and equivalence relation $R_O = \{\leq, \geq, =\}$. C_F and an order relation r compose the ordered structure $\langle C_F, r \rangle$. There are other semantic relations between concepts, such as semantic distance relation, semantic proximity relation and semantic association relation etc.

(4) F is the set of membership functions at U , which is isomorphic to C_F . The corresponding element of F is M in definition 2, but F has also certain structure or relations.

(5) $S = \{s \mid s : C_F \times C_F \rightarrow C_F\}$ is a set of binary operators at C_F . These binary operators form the mechanism of generating new fuzzy concepts. Basic operators are the “union”, “intersection” and “complement” etc., i.e. $S = \{\vee, \wedge, \neg, \Lambda\}$. C_F and S compose the algebra structure $\langle C_F, S \rangle$.

(6) U is the universe of discourse.

Definition 3 is more complex than definition 2 in order to describe the semantic information.

Modeling the linguistic qualifiers, we extend the fuzzy linguistic variable ontology as follows.

Definition 4 (Extended fuzzy ontology) – An extended fuzzy ontology is a 8-tuple $O_F = (c_a, C_F, R, F, S, Q, O, U)$, where:

(1) c_a, C_F, R, F, S, U have same interpretations as defined in definition 3.

(2) Q is the set of the linguistic qualifiers, e.g. $Q = \{\text{very, little, close to, ...}\}$. An qualifier $q \in Q$ and a fuzzy concept $c_F \in C_F$ compose a composition fuzzy concept that can be the value of c_a , e.g. “very cheap”.

(3) O is the set of fuzzy operators at U , which is isomorphic to Q .

To simplify the transform from fuzzy linguistic variables to fuzzy ontology, we introduce the basic fuzzy ontology model as follows.

Definition 5 (Basic fuzzy ontology) – A basic fuzzy ontology is a 4-tuple $O_F = (c_a, C_F, F, U)$, where c_a, C_F, F, U have same interpretations as defined in definition 5, which satisfy the following conditions:

(1) $C_F = \{c_1, c_2, \Lambda, c_n\}$ is a limited set.

(2) Only one relation of set, the relation of disjointness, exists in C_F , and C_F is complete at U . In the other words, C_F is a fuzzy partition of U .

(3) C_F has an ordered relation \leq , and $\langle C_F, \leq \rangle$ is a complete ordered set, i.e. all concepts in C_F constitute a chain $c_1 \leq c_2 \leq \Lambda \leq c_n$.

(4) F is optional element of ontology.

An example of basic fuzzy ontology is $O_F = (c_a = \text{price of product}, C_F = \{\text{very cheap, cheap, appropriate, expensive, very expensive}\}, U = [0, 100])$, where “very cheap” \leq “cheap” \leq “appropriate” \leq “expensive” \leq “very expensive”.

The process to construct fuzzy linguistic variables ontology is as following:

1. Extracting linguistic variables from application domain.
2. Naming the linguistic variables based on domain ontology.
3. Specifying the fuzzy concepts which are the value of fuzzy linguistic variable.
4. (Analyzing the semantic relation between fuzzy concepts.
5. Mapping these fuzzy concepts and their relations to space of membership functions.

For instance, industrial washing machine is used in a variety of domain, but it consumes a mass of water and electricity. Using fuzzy control, the optimal washing strategy which is determined by fuzzy inference can save water and electricity effectively. Fuzzy inference depends on fuzzy knowledge base, i.e. the set of

fuzzy rules, which are summarized from human experiences expressed by fuzzy linguistic variables.

Therefore, definition of fuzzy linguistic variables is the basic of construction of fuzzy rules. To share and reuse fuzzy knowledge, it is necessary to represent formally the fuzzy linguistic variables through ontology. Representing input variables in fuzzy control system for industrial washing machine, there are three basic fuzzy ontologies defined in definition 5 as following:

(1) $O_1 = (c_a = \text{quality of cloth}, C_F = \{\text{chemical fiber, textile, cotton}\}, U = [0,100])$, where “chemical fiber” \leq “textile” \leq “cotton” from point of view of cotton content, and the membership functions are shown in Fig 1.

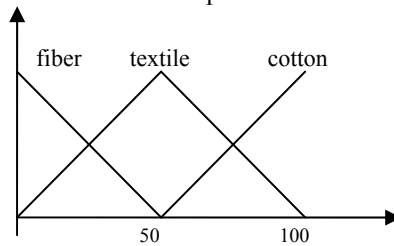


Fig. 1. Membership functions for quality of cloth

(2) $O_2 = (c_a = \text{quantity of cloth}, C_F = \{\text{little, middle, much, very much}\}, U = [0,25])$, where “little” \leq “middle” \leq “much” \leq “very much”, and the membership functions are shown in Fig 2.

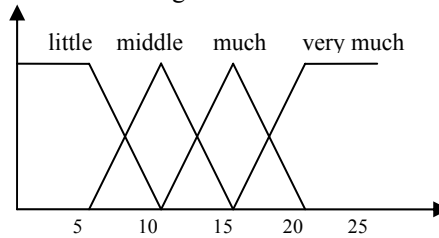


Fig. 2. Membership functions for quantity of cloth

(3) $O_3 = (c_a = \text{degree of squalidity}, C_F = \{\text{clean, dirtish, dirty, filthy}\}, U = [0,100])$, where “clean” \leq “dirtish” \leq “dirty” \leq “filthy”, and the membership functions are shown in Fig 3.

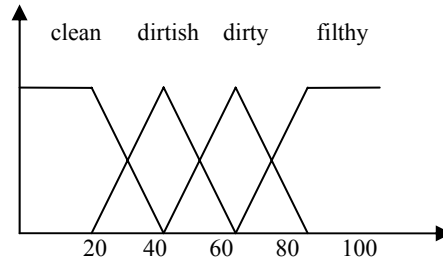


Fig. 3. Membership functions for degree of squalidity

4. Knowledge modeling

In the open and distributed environments of WWW, in order to integrate and reuse information and knowledge in supply chain, ontology becomes the means to model knowledge for customer and product [9]. But the standard ontology is not able to handle fuzzy phenomenon and uncertainty of information and knowledge. It is sufficient for managers and customers to obtain some message in linguistic values rather than in accurate numeric values, such as customer information, product information etc. For instance, the linguistic values for customer income include “low”, “middle”, “high” etc, and linguistic values for product price include “cheap”, “appropriate”, “expensive” etc. These linguistic values have uncertainty and are fuzzy concepts.

The idea of Semantic Web came from Tim Berners-Lee in his vision to move the web into a new generation, where the web resources are annotated with meaning in a form that machines can understand. This will open up vast opportunities for automated processing of the rich knowledge resources available on the web to applications in information search and filtering, knowledge mining, coordination and collaborative processing by intelligent agents. The Semantic Web is to be realized through a shared infrastructure consisting of languages and tools for knowledge representation and processing. The basic knowledge representation format is the Resource Description Framework (RDF) and RDF Schema (RDFS) [10].

Using RDF and RDFS, we construct the ontology structure for customer and product knowledge shown in Fig. 4, in which the linguistic values are represented formally through fuzzy linguistic variable ontologies. The main fuzzy linguistic variable ontologies are as following:

O1=(age, {old, youth, middle-aged, ...});

O2=(income, {little, low, middle, high, ...});

O3=(customer type, {new customer, loyalty customer, gold customer, big customer, lost customer, switched customer ...});

O4=(price, {very cheap, cheap, appropriate, expensive, very expensive});

$O_5 = (\text{zone of influence, \{regional, national, international, \dots\}}); \dots \dots$
 There is a lot of semantic relation between fuzzy concepts. For instance, “gold customer” = “big customer”, “switched customer” \subseteq “lost customer”, “youth” \leq “middle-aged” \leq “old” etc.

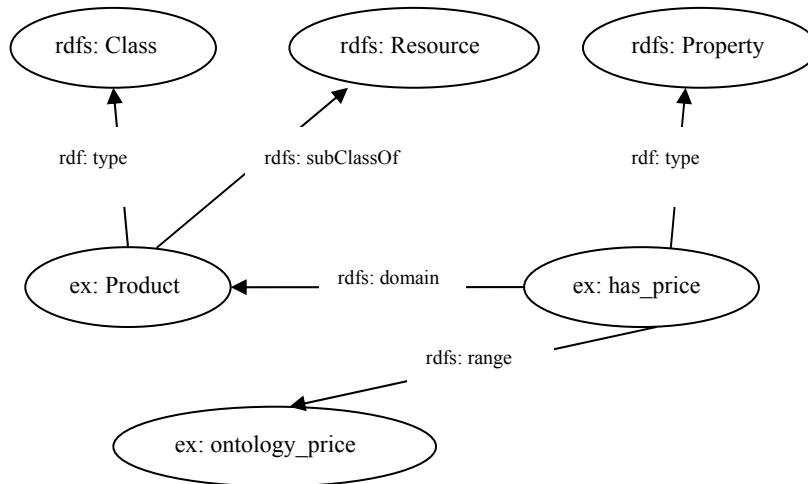


Fig. 4. Ontology structure for product knowledge (portion)

5. Semantic information retrieval

Since the process for information retrieval is based on the knowledge ontology, the semantic and concept research can be achieved. Especially, using linguistic value of fuzzy concept, we can construct the research pattern in SeRQL (Sesame RDF Query Language) [11] such as:

SELECT instance of concept **FROM** {concept} has_property {property} **WHERE** (property of concept) <comparison operator> “Linguistic value of fuzzy concept”, in which the comparison operators includes: equal comparison (=), less than or equal (\leq) and greater than or equal (\geq) etc.

For instance, we can retrieve “product” information through “price” of property, using the search statement such as: **SELECT** Product **FROM** {product} has_price {price} **WHERE** price \leq “expensive”. The standard ontology and other fuzzy ontology are not able to handle the search condition at semantic level, which includes fuzzy concept and semantic relation between them.

Using the “order relation” defined in fuzzy linguistic variable ontology : “very cheap” \leq “cheap” \leq “appropriate” \leq “expensive” \leq “very expensive”, we can transform the search statement to: **SELECT** Product **FROM** {product} has_price

{price} **WHERE** price = “very cheap” or price = “cheap” or price = “appropriate” or price = “expensive”, in which every sub-condition is ordinary and can be completed easily in SeRQL engine.

When retrieving the information about gold customer by the statement: **SELECT** Customer **FROM** {customer} has_type {type} **WHERE** type=“gold customer”, we can obtain the information about big customer using equivalence relation: “gold customer”= “big customer”.

When retrieving the information about lost customer by the statement: **SELECT** Customer **FROM** {customer} has_type {type} **WHERE** type=“lost customer”, we can obtain the information about switched customer using inclusion relation: “switched customer” \subseteq “lost customer”.

Based on other semantic relations defined in fuzzy ontology, we can create and complete more complex semantic retrieval, which will be described in future work.

6. Conclusion

In this paper we have proposed the fuzzy domain ontology model and the fuzzy linguistic variables ontology model to represent fuzzy knowledge. The fuzzy linguistic variables ontology models focus on essential semantic relationships between fuzzy concepts, which facilitates the information retrieval at semantic level. The semantic query expansion in SeRQL query language is constructed by semantic relations between fuzzy concepts.

Our further researches lay on the automatic construction of fuzzy ontology and the integration among standard ontology and fuzzy ontology.

Acknowledgments

This work was supported in part by the Research Project of the Educational Department of Liaoning Province (Leading Laboratory Project) under Grant 20060083.

References

- [1] Fensel D., F. van Harmelen, Horrocks I., D.L.McGuinness, and Patel-Schneider P. F., “OIL: an ontology infrastructure for the semantic web”, IEEE Intelligent Systems, vol. 16 , no. 2, 2001, p. 38-45.

- [2] Widiantoro D. H., Yen J., “A fuzzy ontology-based abstract search engine and its user studies”, in: Proceedings of the 10th IEEE International Conference on Fuzzy Systems, Melbourne, Australia, 2001, p. 1291- 1294.
- [3] Lee C. S., Jian Z. W., and Huang L. K., “A fuzzy ontology and its application to news summarization”, IEEE Transactions on Systems, Man and Cybernetics (Part B), vol. 35, no. 5, 2005, p. 859- 880.
- [4] Tho Q. T., Hui S. C., Fong A. C. M., and Cao T. H., “Automatic fuzzy ontology generation for semantic web”, IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 6, 2006, p. 842- 856.
- [5] Abulaish M., Dey L., “A fuzzy ontology generation framework for handling uncertainties and nonuniformity in domain knowledge description”, in: Proceedings of 2007 International Conference on Computing: Theory and Applications, Kolkata, 2007, p. 287-293.
- [6] Kang D. Z., Xu B. W., Lu J. J., Li Y. H., “Description logics for fuzzy ontologies on semantic web”, Journal of Southeast University (English Edition), vol. 22, no. 3, 2006, p. 343 – 347.
- [7] Calegari S., Ciucci D., “Fuzzy ontology and fuzzy-OWL in the KAON project”, in: Proceedings of 2007 IEEE International Conference on Fuzzy Systems Conference, London, UK, 2007, p.1-6.
- [8] Jun Zhai, Yan Chen, Qinglian Wang, and Miao Lv, “Fuzzy Ontology Models Using Intuitionistic Fuzzy Set for Knowledge Sharing on the Semantic Web”, in: Proceedings of the 12th International Conference on Computer Supported Cooperative Work in Design (CSCWD 2008)(volume 1), 2008, IEEE Press, p.465-469.
- [9] Lu X. W., Jiang F., Hou L. W., “Customer features extraction based on customer ontology”, Computer Engineering, vol. 31, no. 5, 2005, p. 31-33. (in Chinese)
- [10] John W.T. Lee, Alex K.S. Wong, “Information retrieval based on semantic query on RDF annotated resources”, in Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics, 2004, p. 3220-3225.
- [11] .B. V. Aduna, “The SeRQL query language,”
<http://www.openrdf.org/doc/sesame/users/ch06.html#d0e1977>, 2002.