

Ontology Mapping based on Similarity Measure and Fuzzy Logic

S. NIWATTANAKUL¹

Ph. MARTIN²

M. EBOUEYA¹

K. KHAIMOOK³

¹ Informatics, Image and Interaction Research Laboratory, University of La Rochelle, France

² School of Information and Communication Technology, Griffith University, Australia

³ School of Information Technology, Suranaree University of Technology, Thailand

suphakit.niwattanakul@univ-lr.fr phmartin@phmartin.info michel.eboueya@univ-lr.fr kkanit@sut.ac.th

Abstract: In this paper, we present a method of an ontology mapping based on a similarity measure and Fuzzy logic in order to classify (i) the similarity of the ontology structure of learning object repositories and (ii) LOR which stores metadata of learning objects based on our ontology model. In this model, values of the ontology similarity are computed for concepts, properties, and relations. The ontology similarity uses parameters based on the Fuzzy Control Language (FCL) which consists of a fuzzy set of the ontology similarity (“Less”, “Same”, “More”), 7 classes of ontology similarity, and rules of the classification of ontologies. The formula of similarity measure by the Jaccard’s coefficient is applied to map a similarity of ontology structures. At the end of the article, we show an experience of implementation this model as a prototype.

Introduction

Nowadays, ontologies are applied to various domains such as e-learning where they are used for describing metadata of learning resources to help searching and retrieving learning objects in their repository. However the usage of ontology structures may be different in some learning object repositories and this leads to the problem of searching learning objects in various learning object repositories. To solve this problem, we propose a model to define and classify an ontology similarity, which is meant the similarity of ontology structures, in order to select suitable learning object repositories for the searching and retrieving learning objects in these repositories.

Our Learning Object Repository (LOR) is a system that stores learning objects on the Web and/or their metadata for the serving of searching and retrieving learning objects on the Internet and the structure of metadata is based on ontology model.

In this paper, we apply a similarity measure for ontology mapping in order to compute the probability of the similarity of ontology structure and Fuzzy Logic is used for classifying the ontology similarity. The technique of ontology mapping in our system is implemented by (1) the comparison of 3 main categories of ontology structure: concepts, properties, and relations, (2) the computation of the probability of the similarity measure, and (3) the classification of ontology similarity with Fuzzy Logic.

Definitions and purposes of an ontology

Gruber (1993) defines an ontology as “*a specification of a conceptualization*”. Another definition is given in (Studer et al, 1998): “*An ontology is a formal, explicit specification of a shared conceptualization*”. The specification of the conceptualization consists of the objects, concepts and other entities that are in the same particular domain and the relationships that hold among them. “Explicit” means that objects, concepts, and other entities are explicitly defined. “Formal” implies that the ontology should be machine readable. “Shared” means that the ontology captures consensual knowledge and is agreed-upon by a group, not just an individual. In generally, the main structure of ontology model consists of tree main categories: concept or class, property, and relation.

In learning resource management, ontology has been applied to improve the structure and the usefulness of Learning Design Repository in IDLD (Implement and Development of the Learning Design) project (see Paquette et al, 2006). The use of the ontology could be employed as an approach to implement a semantic Web-based e-learning system. This framework is focused on the RDF (Resource Description Framework) data model, OWL ontology language and RAP for parsing RDF documents (Fayed et al, 2006).

Recently, research on the ontology technology aims at the act of interoperability and reusability. It means that similar objects which are described in different ontology structures could be integrated into a new ontology structure and they could be utilized in a particular system. This technology is known as an ontology mapping. As described in (Laurel et al, 2004), there are two types of ontology mapping: source-based and instance-based. Examples of source-based mapping tools are PROMPT, Chimaera, and ONION and examples of instance-based

mapping tools are FCA-Merge and GLUE. Beyond, a new methodology for merging the heterogeneous domain ontologies based on the WordNet which is used as a dictionary to give relationships between concepts detailed in (Kong et al, 2005).

Defining and classifying the ontology similarity

In this section, we present how to define and classify the ontology similarity. The manner of the similarity measure is used for defining the ontology similarity while the Fuzzy Logic model is applied to classify the ontology similarity.

Similarity measure of ontologies

Similarity measure can define a similarity of any two ontologies involved. The well-known formula of similarity measure is Jaccard's coefficient which appears in (Doan et al, 2002) as shown following:

$$\begin{aligned} \text{Jaccard's } sim(A,B) &= P(A \cap B) / P(A \cup B) \\ &= \frac{P(A,B)}{P(A,B) + P(A, \overline{B}) + P(\overline{A}, B)} \end{aligned} \quad (1)$$

To detail the formula (1), we assume that A and B are any ontology structure. $P(A, B)$ is a number of elements of the ontology A and B which are similar. $P(A, \overline{B})$ is a number of elements of the ontology A which are similar with elements of the ontology B. $P(\overline{A}, B)$ is a number of elements of the ontology B which are similar with elements of the ontology A.

The Jaccard's coefficient is practically applied to define similarity measure in many systems such as the GLUE system which employs this formula for machine learning techniques to find mapping in the ontology model. This system will map the similarity of two ontologies into a new ontology (Doan et al, 2002). A methodology of ontology mapping proposed by Laurel et al (2004) is based on similarity measurement by Jaccard's coefficient as well.

For computing the probability of the similarity measure of ontology structure in each mediator, we begin to compare the structure of ontology in three categories: concepts, properties, and relations. Then values of the probability of ontology similarity each category are computed by using Jaccard's coefficient. Values of the probability will be between 0 and 1. The methodology is shown by the algorithm proposed as follows:

```
// Defining the ontology similarity
For each mediator // outer loop
    Read concepts, properties, and relations for outer mediator
    For each other mediators // inner loop
        Read concept, properties, and relations or inner mediator
        // For concept category
            Compare all concepts between outer and inner mediator
            Compute value of the probability of similarity of concept
        // For property category
            Compare all properties between outer and inner mediator
            Compute value of the probability of similarity of property
        // For relation category
            Compare all relations between outer and inner mediator
            Compute the value of the probability of similarity of relations
    Next inner loop
Next outer loop
```

Classification the similarity of ontology with Fuzzy Logic

Basically, Fuzzy Logic (FL) is a multivalued logic that allows intermediate values to be defined between conventional evaluations like yes/no, true/false, black/white, etc. The basic notion of fuzzy systems is a fuzzy set, for example to classify the fuzzy set of climate which may be consisted of member like "Very cold", "Cold", "Warm", "Hot", and "Very hot".

Similarity measure based on Fuzzy Logic is proposed in (Idri and Abran, 2001) to measure for software project similarity. And it has also been applied to analyze the similarity of dynamical system with voice processing (Rodriguez, 2000). We now propose the methodology of classification of similarity of ontology in the way of Fuzzy Logic. Firstly, the constraints of concepts, properties, and relations and rules are based upon Fuzzy Control Language (FCL) which is a language for implementing Fuzzy Logic standardized by IEC 1131-7¹. Secondly, all values of the probability of the ontology similarity from prior section are classified in similarity of ontology structure. Finally, the results of classification of the ontology similarity classified by rules from FCL will be presented. To classify the similarity of ontology structure in our model, we propose to define the definition of similarity of ontology into three levels of fuzzy set and three constraints.

In our design, three categories consist of concepts, properties, and relations and a fuzzy set of ontology similarity of each category contains “Less”, “Same”, and “More”. The range of the fuzzy set is from 0 to 1 and the similarity of ontology is defined by the value of the probability of the ontology similarity as shown in Figure 1.

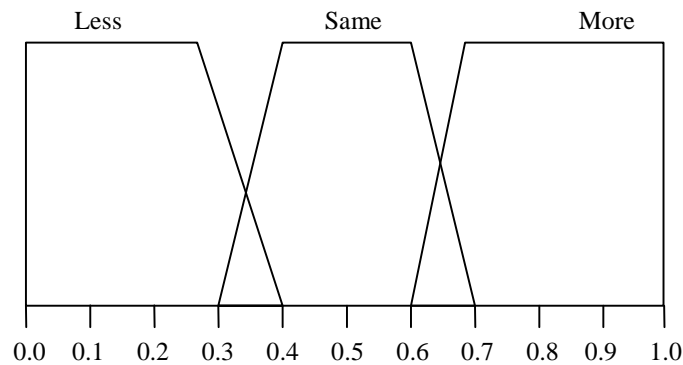


Fig 1. Range of the fuzzy set for defining of the ontology similarity

The similarity of ontology is divided into 7 classes which are “Class_0”, “Class_1”, “Class_2”, “Class_3”, “Class_4”, “Class_5”, and “Class_6”. “Class_0” is the least similar while “Class_6” is the most similar. The range of the classification of the ontology similarity is from 0 to 1 and the classification of the ontology similarity is defined by value of the classified similarity of ontology structure as shown in Figure 2.

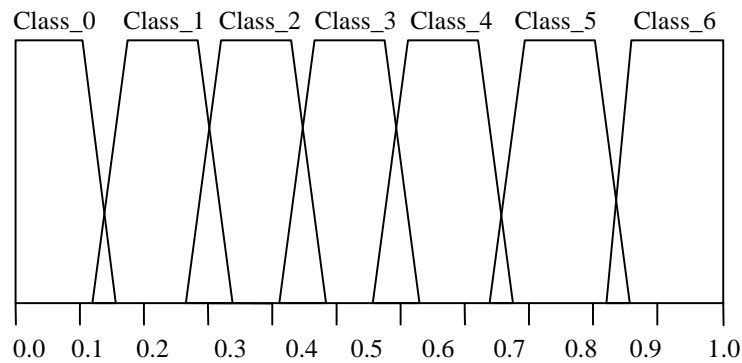


Fig 2. Range of the fuzzy set of the classification of the ontology similarity

To classify the similarity of ontology structure, the fuzzy rule set is designed in Fuzzy Control Language shown in the following examples:

- RULE 1 : IF (concept IS More) and (property IS More) and (relation IS More) THEN similarity of ontology IS Class_6;
- RULE 2 : IF (concept IS More) and (property IS More) THEN similarity of ontology IS Class_6;

¹ International Electrotechnical Commission (IEC). IEC 1131 – Programmable Controllers: part 7 Fuzzy Control Programming. 1997

RULE 3 : IF (concept IS More) and (relation IS More) THEN similarity of ontology IS Class_6;
 RULE 4 : IF (property IS More) and (relation IS More) THEN similarity of ontology IS Class_6;

RULE 5 : IF (concept IS More) and (property IS More) and (relation IS Same) THEN similarity of ontology IS Class_5;
 RULE 6 : IF (concept IS More) and (relation IS Same) THEN similarity of ontology IS Class_5;
 RULE 7 : IF (property IS More) and (relation IS Same) THEN similarity of ontology IS Class_5;
 ...

From previous steps, the fuzzy set for defining of the ontology similarity and fuzzy rule set are designed in the format of Fuzzy Control Language to FCL file. The FCL file will be read by Fuzzy Logic tool for the classification of the ontology similarity. The function of classification of similarity of ontology is implemented as an algorithm as follows:

```
// Classifying of similarity of ontology
Read Fuzzy Logic rules from a file which is in the format of Fuzzy Control Language (as described above)
For each mediator // outer loop
  For each other mediator // inner loop
    Set values to concept, property, and relation variables
    Classifying with Fuzzy Logic rules
  Next
Next
Show results
```

Experimentation

To experiment with algorithm described in the previous section, we have tested these algorithms with our prototype, coding by JavaServer Pages Technology (JSP)² and using MySQL³ as a database tool. We have started with 10 simulated mediators that each mediator consists of 20 concepts, 20 properties, and 20 relations with random names. Then we compute the similarity of concepts, properties, and relations each mediator based on similarity measurement by Jaccard's coefficient. The similarity of each mediator will be classified by the Open Source library of Fuzzy Logic tool using jFuzzyLogic⁴. As shown in Table 1, the values of classification of similarity of mediators are shown as the result of our prototype. For example, when we search learning objects in mediator "M_01" and the mediator "M_01" want to search other learning objects in other mediator. The values from Table 1 can be used to decide where it should start searching learning objects. In this example, it should start to search in mediator "M_03" and then "M_07", "M_09", and so on respectively. We believe that this model can help to improve an efficiency of the search learning objects in multi mediators by selecting suitable mediator.

Table 1: The values of classification of similarity of mediators

Mediator	M_01	M_02	M_03	M_04	M_05	M_06	M_07	M_08	M_09	M_10
M_01	+	0.5500	0.7614	0.3878	0.6396	0.5000	0.7157	0.5000	0.7034	0.5517
M_02	0.5000	+	0.6122	0.5000	0.5000	0.5000	0.5000	0.6396	0.6122	0.5000
M_03	0.7614	0.6122	+	0.5000	0.8742	0.5000	0.9258	0.5000	0.9258	0.5000
M_04	0.3878	0.5000	0.5000	+	0.5000	0.5000	0.5000	0.5000	0.3878	0.5000
M_05	0.6396	0.5000	0.8742	0.5000	+	0.5000	0.7614	0.5000	0.7034	0.5000
M_06	0.5000	0.5000	0.5000	0.5000	0.5000	+	0.5517	0.5000	0.5000	0.5000
M_07	0.7157	0.5000	0.9258	0.5000	0.7614	0.5517	+	0.5000	0.9258	0.5000
M_08	0.5000	0.6393	0.5000	0.5000	0.5000	0.5000	0.5000	+	0.5000	0.5000
M_09	0.7034	0.6122	0.9258	0.3878	0.7034	0.5000	0.9258	0.5000	+	0.5000
M_10	0.5517	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	+

² JavaServer Pages Technology. <http://java.sun.com/products/jsp/>

³ MySQL. <http://www.mysql.com/>

⁴ jFuzzyLogic: Open Source Fuzzy Logic (Java). <http://14logic.sourceforge.net/html/index.html>

The advantages of defining and classifying the similarity of ontology structure in our system are: (1) the implementation is based on the ontology mapping technique and Fuzzy Logic, (2) The similarity of mediator can be classified the similarity by using the similarity measure of Jaccard's coefficient and in the way of Fuzzy Logic.

Conclusions

In this paper, we have presented the defining and classifying the similarity of ontology structure. To classify the ontology similarity, the similarity measure of ontology will be computed from three categories; concepts, properties, and relations by using Jaccard's coefficient. The Fuzzy Control Language is applied to classify similarity of ontology by designing the fuzzy set which consists of three levels: "Less"; "Same"; and "More" and classification of resources in 7 classes.

In our experience, we have found that the classification of mediator similarity can be effect to retrieve learning objects in various learning object repositories by searching only similar mediator. In the future, we could extend our work to experiment on classifying the similarity of learning objects in our model.

References

- Doan, J. Madhavan, P. Domingos and A. Halevy. (2002) **Learning to map between ontologies on the semantic web**. In Proc of the 11th International WWW Conference.
- Fayed F. M. Ghaleb, Sameh S. Daoud, Ahamad M. Hasna, Jihad M. Jaam, and Hosam F. El-Sofany. (2006) **A Web-Based E-Learning System Using Semantic Web Framework**. *Journal of Computer Science* 2 (8), pp. 619-626),
- Gruber, T.R. (1993) **A translation approach to portable ontologies**. *Knowledge Acquisition*. [Online]. Available: http://ksl-web.stanford.edu/KSL_Abstracts/KSL-92-71.html
- Idri, A. Alain Abran, A. (2001) **A Fuzzy Logic Based Set of Measures for Software Project Similarity: Validation and Possible Improvements**. *IEEE METRICS 2001*: pp. 85-97, 2001.
- Kong, H. Hwang, M. and Kim. P. (2005) **A New Methodology for Merging the Heterogeneous Domain Ontologies based on the WordNet**. In Proc of *The International Conference on Next Generation Web Services Practices (NWeSP'05)*.
- Laurel C. Y. Kong, C. L. Wang, and F. C. M. Lau. (2004) **Ontology Mapping in Pervasive Computing Environment** *International Conference on Embedded and Ubiquitous Computing (EUC-04)*. pp. 1014-1023. Aizu, Japan, 26-28 August 2004.
- McGuinness, D. L. and Harmelen, F. V. (2004) **OWL Web Ontology Language Overview, W3C**. [Online]. Available: <http://www.w3.org/TR/owl-features/>.
- Paquette, G. Marino, O. Lundgren-Cayrol, K. Léonard, M. Teja, I. de la. (2006) **Learning Design Repositories – Structure Ontology and Processes**. In Proc. of *International Workshop of Learning Networks for Lifelong Competence Development*. pp. 18-22. March 30-31, 2006. Sofia, Bulgaria.
- Rodriguez W. (2000). **Fuzzy Logic Based Voice Processing**. In Proc. of *The International conference on: Modeling, Simulation and Neural Network (MSNN-2000)*. October 22-24, 2000. Mérida, Venezuela, 2000. [Online] Available: <http://iies.faces.ula.ve/Amse2000/papers/fuzzy/MSNN2000Wladimir.pdf>
- Studer, R. Benjamins, V.R. and Fensel D. (1998) **Knowledge Engineering: Principles and Methods**. *Data and Knowledge Engineering*, vol.25. pp. 161-197.