Chapter 126
Ontology-Based Multi-Enterprise Heterogeneous Model Fusion Method

Hong-xiu Wang

Abstract For the multi-enterprise collaborative modeling environment, the semantic conflicts of concept in the merger from the local model into the overall model or in the integration from the lower model into the upper model, the semantic-based multi-enterprise heterogeneous model fusion method is proposed. Moreover, the semantic similarity among the model instances is analyzed from the various levels, and based on the semantic similarity, a series of rules of the model merging are proposed, and then model integration is completed. Finally, the similarity matching tool is developed to realize based on semantic similarity analysis.

Keywords Heterogeneous · Model integration · Multi-enterprise model · Ontology

126.1 Introduction

In the collaborative enterprise modeling, the model is completed by more than one person in the project team; they will apply their own terms to create a model instance, resulting in the semantic conflict in the merger from the partial model to the overall model. The main problems are: the one on the same physical application of different terms to describe, the same terminology to describe different content, three different definitions of the granularity of the process, activity.

For this type of semantic heterogeneity, the related research work focus on building a unified dictionary based on meta-data (Castano et al. 2005; Missikof and Schiappelli 2003), but less involved in the essence of the information

H. Wang (✉)
Department of Industrial Engineering, Tianjin Polytechnic University, Tianjin, China
e-mail: littlestone25@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI: 10.1007/978-3-642-38391-5_126,

semantics, and can not fundamentally solve the problem (Cui et al. 2001; Mike et al. 1998). On the basis of the work on the research at home and abroad, proceed from solving the semantic heterogeneity of shared modeling the collaborative modeling, proposed an enterprise ontology-based concept of constraints to solve the consistency problem in enterprise modeling.

Ontology achieves the effective semantic understanding and communication between people or application systems (Horrocks et al. 2003; Pulido et al. 2006). In engineering applications, the ontology can support semantic interoperability. It provides the mechanism described in the explanation of the objective world. Semantic interoperability requirements of the data easier to understand, and can be easily defined mapping between the data of known and unknown (Athena 2004; Berre et al. 2007).

126.2 Multi-Model Fusion Method

126.2.1 Two Assumptions

Generally, the enterprise model is composed of multiple views, and its structure is complex. In order to achieve the integration of partial models, first of all assume that the two conditions have been met: First, the model is divided into view established; Second, during the merger have been identified belong to the same view of part of the model can be combined with a view of the upper model.

126.2.2 The Formal Definition of the Enterprise Model

According to the references (Vernadat 2002) on the definition of enterprise model and (Thiagarajan et al. 2006) on the definition of the business model, it is present that the business model formal definition is as follows:

Definition 1 Enterprise model EM = (E, R), where E is composed of a collection of elements for the model, R is the relationship between these elements in the model.

Modeling element E = {{Attribute}, {range}, {Subelem}}, Attribute is description of the properties of the element characteristics, range is range of the property, Subelem is set of elements of sub-concepts.

126.2.3 Similarity Definition

Definition 2 The formal definition of the similarity of two concepts x, y :
Sim(x,y)∈[0…1]
Sim(x,y) = 1→x = y: The two entities are equivalent
Sim(x,y) = 0 : The two entities do not intersect, unrelated
Sim(x,y) = Sim(y,x) : the symmetry

Matching relations between concepts, we can make a judgment according to their similarity.

In this paper, Sim, (c1, c2) expressed the similarity function between the two concepts c1 and c2.

Set a threshold t, when Sim (c1, c2) ≥ t, we believe that c1 and c2 is similar.

126.2.4 Analysis of Ontology-Based Model Semantic Similarity

Basis for the formal definition of the enterprise model, in order to calculate comprehensively, accurately the similarity between the concepts, Respectively, it is calculated based on the name, concept properties, the subset of concept. Finally, it is given the right value to merge the similarity.

(1) Calculation of concept name similarity. Assumptions two concepts A and B, the similarity of their names is calculated as:

$$sim_{name}(\mathrm{A}_{name}, \mathrm{B}_{name}) = \frac{N(\text{the longest substring between } \mathrm{A}_{name} \text{ and } \mathrm{B}_{name})}{(N(\mathrm{A}_{name}) + N(\mathrm{A}_{name}))} \tag{126.1}$$

If the concept has an alias, in addition does computing the concept name similarity, but also to calculating the similarity of the concept alias. Using the formula (126.1), the final name similarity is

$$Sim_{nameZ} = \sum_{j=1}^{m+1} \sum_{i=1}^{n+1} w_{ij} Sim_{name}(A_i, B_j) \ \mathrm{n, m} \geq 0 \tag{126.2}$$

Among them, $\sum_{j=1}^{m+1} \sum_{i=1}^{n+1} w_{ij} = 1$, n is the number of the alias of the concept A, m is the number of the alias of the concept B. When n = 0, m = 0, $Sim_{nameZ} = Sim_{name}$

(2) Calculation of similarity based on concept attributes

Based on the attributes, the theoretical basis for calculation of the conceptual similarity is: if the attributes of the two concepts are the same, then the two concepts are the same; if the two concepts have similar properties, these two concepts are similar. Each concept in the ontology is to be described and limited by a set of attributes. The attribute set definition is given in the following.

Definition 3 Let A = $\{A_1[V_1], A_2[V_2], \ldots, An[Vn]\}$, A is a set of properties. A_i is the attribute name; V_i is the range of A_i. The definition is the set of attributes were classified into the attribute set level and the attribute value level. The calculation of similarity of the property is divided into two parts of the set of attributes and attributes values to conduct investigations. Let C1 and C2 are the concept associated attribute set of the objects o_1 and o_2. The similarity of the attribute set is:

$$Sim_{attrS} = \frac{1}{|dist(o_1, o_2) - 1|} \times \frac{|C_1 \cap C_2|}{|C_1 \cap C_2| + \alpha|C_1 - C_2| + (1 - \alpha)|C_2 - C_1|} \tag{126.3}$$

There may be different values in the instances of two objects in the common property. Therefore, the value of the similarities and differences in the common property need inspect. Let $A_i|C_1 \cap C_2|$. $A_i(o)[v]$ represents that the value of the instance o on attribute A_i is v, and the upper and lower bounds of the statistical range of the A_i values are expressed as Low(A_i), High(A_i). The similarity of the attribute value is:

$$Sim_{attrV} = \prod_{i=1}^{|C_1 \cap C_2|} \left(1 - \frac{|A_i(o_1)[v_1] - A_i(o_2)[v_2]|}{|Low(A_i) - High(A_i) + 1|}\right) \tag{126.4}$$

According to Ai specific data types, the specific meaning of its statistical range is different. For example, for the numerical data type, the difference between the maximum and minimum can be used in the actual value of the attribute. For Boolean data type, 0, 1 value is processed. For string type, if the attribute values of two instances are the same, similarity is 1, otherwise 0.

In the end, the similarity of two instances in the characteristics of the attribute set is the superposition of these two aspects. The formula is

$$Sim_{attribute} = Sim_{attrS} \times Sim_{attrV} \tag{126.5}$$

In addition, a concept may have multiple attributes and the effects and the extent described of each attribute on the concept are different. Therefore, if each attribute is involved, the amount of the calculation will be greatly increased. When the attribute similarity is calculated, the attributes need be classified, and focusing on the calculation of the business attributes.

(3) Calculation of similarity based on a set of concept

In the ontology, the meaning of a concept can consist of the meaning of its direct sub-concepts. The combination of all sub-concepts can describe the meaning of the concept. Thus, the similarity between the upper concepts can be obtained by calculating the similarity between the sub-concepts. This method is flexible and extensible. Let A, B for the two upper concept in the ontology, similarity between A and B using following formula:

$$Sim_{sub}(A,B) = \frac{\sum_{a_i \in A} \max_{b_j \in B} S(a_i, b_j) + \sum_{b_j \in B} \max_{a_i \in A} S(b_j, a_i)}{N(A) + N(B)} \quad (126.6)$$

N(A) indicates the number of sub-concepts of A, N(B) indicates the number of sub-concepts of B. S(a,b) is calculated by using the instance-based method, formulated as:

$$Sim(A,B) = \frac{P(A \cap B)}{P(A \cup B)} = \frac{P(A,B)}{P(A,B) + P(\overline{A},B) + P(A,\overline{B})} \quad (126.7)$$

which P(A,B) indicates the probability that this concept is sub-concepts both A and B when a concept randomly is selected from the ontology.

$$P(A,B) = \left(N(U_1^{A,B}) + N(U_2^{A,B})\right) \Big/ \left(N(U_1) + N(U_2)\right) \quad (126.8)$$

U_i indicates the set of underlying concepts in the ontology i, $N(U_i)$ indicates the number of the concepts in U_i. $N(U_1^{A,B})$ indicates the number of the concepts both belong A and B in the ontology 1. $N(U_2^{A,B})$ indicates he number of the concepts both belong A and B in the ontology 2. At this point, the similarity of A and B is obtained.

(4) Comprehensive computation of similarity

This three kinds of similarities are comprehensively computed, the formula of the final comprehensive similarity as follows:

$$Sim(A,B) = w_{name}Sim_{nameZ}(A,B) + w_{attribute}Sim_{attribute}(A,B) + w_{sub}Sim_{sub}(A,B)$$

which, $w_{name} + w_{attribute} + w_{sub} = 1$.

126.2.5 E Model Merging Rules Based on Semantic Similarity

Setting a threshold for the above four kinds of similarity, the threshold is usually determined by experts or analysts. When the calculated similarity is greater than the threshold, they are called name similarity, attribute similarity, subset similarity and comprehension similarity. Where, the model merging rules are defined based on these four similarity relations. According to these rules, and then the overall model is generated.

Rule 1: if the two model instances are comprehension similarities, the one is kept, another is deleted in the model merging.

Rule 2: if the two model instances are name similarities and the similarity is less than 1, but the attribute and the subset are not similar, two models are kept in the model merging.

Rule 3: if the name similarity of the two model instances is equal to 1, but the attribute and the subset are not similar, two models are kept in the model merging. At the same time, the name of a model is modified.

Rule 4: if two models are name similarity and attribute similarity, but their subsets are not similar, two models are kept in the model merging.

Rule 5: if two models are name similarity and subset similarity, but their attributes are not similar, two models are kept in the model merging.

Rule 6: if two models are attribute similarity and subset similarity, but their names are similar, one model is kept in the model merging.

Rule 7: if the two models are attribute similarity, but their names and subsets are not similar, two models are kept in the model merging.

Rule 8: if the two models are subset similarity, but their names and attributes are not similar, two models are kept in the model merging.

Rule 9: if the two models are not subset similarity, name similarity, attribute similarity, two models are kept in the model merging.

126.3 The Tool System of Model Knowledge Matching

The function of the system of model knowledge matching consists of ontology editing, database/owl transformation and concept matching. The ontology editing module has completed the editing and maintenance functions of ontology, establishing ontology tree and all knowledge stored in the database. Database/owl transformation module has completed to transformation the knowledge stored in

Fig. 126.1 Concept matching

the database into the standard expressed in OWL ontology. Concept matching module has completed to compute the similarity of the inputted concept.

According to the precious method of similarity calculation, when two concepts and the weight of each similarity input, the multi-layer similarities and the total matching are computed. Figure 126.1 show the matching result between "Quotation" and "Payment application form".

126.4 Conclusion

In this paper, the model merging method from the partial model to the whole model is studied. Moreover, the semantic similarity among the model instances is analyzed from the various levels, and based on the semantic similarity, a series of rules of the model merging are proposed, and then model integration is completed.

Finally, a prototype system of the model knowledge matching is developed. And a case is described to validate the modeling method proposed in this paper.

References

Athena (2004) Advanced technologies for interoperability of heterogeneous enterprises networks and their applications. FP6-2002-IST-1, integrated project description of Work, pp 19–36

Berre A, Elvesæter B, Figay N, et al (2007) The ATHENA interoperability framework. In: Proceedings of the 3rd international conference on interoperability for enterprise software and applications, enterprise interoperability II. Springer, Madeira, pp 569–580

Castano S, De Antonellis V, De Capitani di Vimercati S (2005) Global viewing of heterogeneous data sources. IEEE Trans knowl Data Eng 13(2):277–297

Cui Z, Jones D, O'Brien P (2001). Issues in ontology-based information integration. http://www.csd.abdn.ac.uk/ebiweb/papers/cui.pdf

Horrocks I, Peter F, Van Harmelen Frank (2003) From SHIQ and RDF to OWL: the making of a web ontology language. J Web Seman 1(1):7–26

Mike U, Martin K, Stuart M (1998) Enterprise ontology. Knowl Eng Rev 13(1):31–98

Missikof M, Schiappelli F (2003) A controlled language for semantic annotation and interoperability in e-business applications. In: Proceedings of the second international semantic web conference, Sanibel Island, pp 1206–1211

Pulido JRG, Ruiz MAG, Herrera R et al (2006) Ontology language for the semantic web: a never completely updated review. Knowl Based Syst 19(7):489–497

Thiagarajan RK, Srivastava AK, Pujari AK (2006) BPML: a process modeling language for dynamic business models. In: Proceedings of the 4th IEEE international workshop on advanced issues of e-commerce and web-based information systems, Califoria, pp 222–225

Vernadat F (2002) UEML: towards a unified enterprise modeling language. Int J Prod Res 40(17):4309–4321