Semantic Inconsistency Errors in Ontology

[1]Muhammad Fahad, [2]Muhammad Abdul Qadir, [3]Muhammad Wajahaat Noshairwan
Mohammad Ali Jinnah University, Islamabad, Pakistan
[1]mhd.fahad@gmail.com, [2]aqadir@jinnah.edu.pk, [3]wajahaat@jinnah.edu.pk

Abstract

Ontology evaluation is one of the most important phases of Ontology Engineering. Researchers have identified different types of errors that should be catered in ontology evaluation process and classified them in error's taxonomy. Semantic Inconsistency Errors are the most common when modeling complex taxonomic knowledge while building semantic classification within the ontology, and in ontology merging process where different subtype concepts of source ontologies map on each other. In this paper, we have identified three types of semantic inconsistency errors, and provided criteria for semantic classification evaluation. This criterion provides ontologists to build well-formed class hierarchy that is free from semantic inconsistency errors. Moreover we emphasized on the control mechanisms (based on this criteria) to be perform during finding mappings between concepts in ontology merging process to achieve accurate results. We have demonstrated the importance of such errors by giving different scenarios where appropriate.

1. Introduction

Ontology is regarded as the formal specification of the knowledge of concepts and the relationships among them [3]. They require formal syntax and semantics to represent domain concepts. They have played a key role for describing semantics of data in both traditional knowledge engineering and the emerging Semantic Web [12]. Ontology has to go through a repetitive process of refinement during its development lifecycle. Ontology engineers have to pay much attention to produce high quality bug free ontology. But there is a possibility that the ontologists unintentionally make some errors in ontology [5]. Ontology evaluation is one of the most important phases of Ontology Engineering because if ontology itself is error prone then the applications dependent on the ontology have to face some critical and catastrophic problems.

Domain researchers have identified some errors and defined them in error's taxonomy for assistance in the ontology evaluation [1, 9]. This error's taxonomy becomes a guideline for ontology engineer to evaluate the ontology in perspective of such errors. Gomez et al. identified and categorized three types of errors that are usually encountered by ontologist i.e. inconsistency, incompleteness and redundancy of information. Inconsistency means the ontology have any type of contradictory information [3]. Redundancy means that same information is inferred from ontology more than once. Incompleteness means the concepts are not completely defined. Qadir et al. identified the scenarios where incomplete partition error such as disjoint knowledge omission leads toward catastrophic situations [4]. They gave the criteria to detect concepts of a given ontology that requires disjoint knowledge between concepts. Baumeister and Seipel [6] discussed the design anomalies in ontology and defined the detection method by using prolog and FN-query language. These anomalies help the ontologist to develop consistent ontology. Brank et. al. [8] discussed the overall approaches for ontology evaluation and concludes that different approaches are useful in different application. Previously we have extended Gomez's taxonomy of errors by identifying two new errors [9]. One is Sufficient Knowledge Omission Error that arises when ontologists does not elaborate the characteristics of the concept like its self description by using intersection, union, complement or restriction axioms in OWL. Second is Redundancy of Disjoint Relation Error that arises when ontologists made concept disjoint with other concept more than once.

In this paper, we present three types of semantic inconsistency errors that can be made by ontologists when modeling taxonomic knowledge in ontology. These semantic errors are increasingly seen while finding correspondences between concepts of source

0-7695-3032-X/07 $25.00 © 2007 IEEE
DOI 10.1109/GrC.2007.153

ontologies during developing our semantic based ontology merger, DKP:OM [11]. Ontology merging is an active research area that answers the questions about the combined use of independent developed ontologies [13]. Different ontology merging systems are developed that enable us to integrate different source ontologies and allow interoperability among them. These systems find the places in the ontologies where concepts overlap, relate concepts that are semantically close by equivalence and subsumption relations [12]. We face semantic inconsistency errors when DKP:OM finds partial mappings between concepts that lead to an incorrect semantic classification, that is, it produces a mapping of a concept as a subclass of another concept to which it does not really belong. For semantic classification evaluation, we identified three types of semantic inconsistency errors. We have defined these errors and the situations where they can occur and explained the importance of these errors by different examples or scenarios.

Rest of the paper is organized as follows: section 2 elaborates semantic inconsistency errors; section 3 presents semantic inconsistency in ontology merging process. Section 4 concludes the paper and gives insight on future work.

2. Extension in Error's Taxonomy

Figure 1 shows the error's taxonomy by Gomez-Perez [1], slightly extended by us. The contribution of this paper shows the categorization of semantic inconsistency errors that is represented in dotted box.

According to Gomez, these errors occur when ontologists make an incorrect class hierarchy by classifying a concept as a subclass of a concept to which it does not really belong [3]; for example he classifies the concept Airbus as a subclass of the

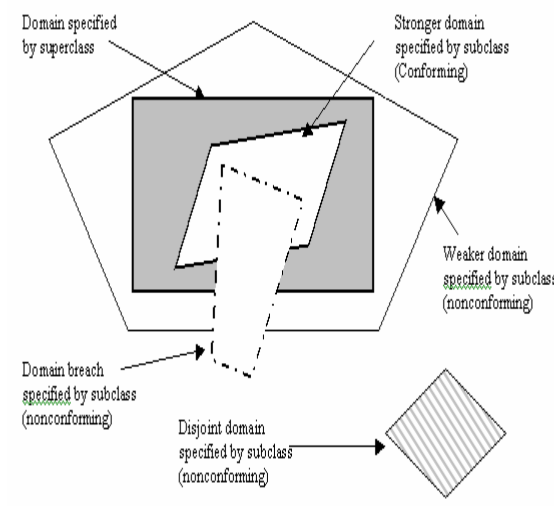

Figure 2. Conformance relationship between superclass and subclass

concept Train. Or the same might did when classifying instances. There are mainly three reasons due to which this type of incorrect semantic classification originates. According to these reasons we categorize classification Inconsistency errors into three subclasses. These subclasses can be used as a check list for classification evaluation and help in building well-formed class hierarchy to provide better interpretation of concepts. Figure 2 shows the conformance relationships between superclass and subclass concepts. These may be applied same to the instances of superclass and subclass as well.

2.1. Weaker domain specified by subclass error

These errors occur when ontologists specify subclass that possesses weaker domain as compared to the superclass domain. Usually this type of error occurs when building class hierarchies in ontology and

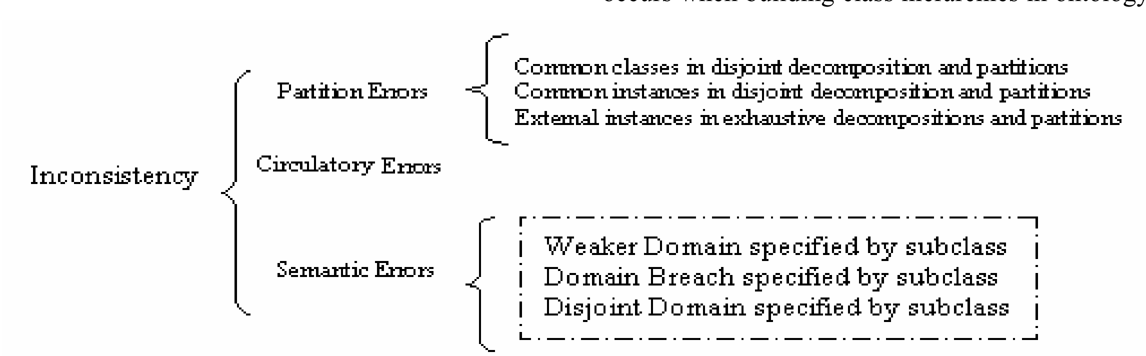

Figure 1. Extended Error's Taxonomy

For example he classifies accommodation, House, place, and Shelter concepts as a subclass of a concept Hotel superclass. This classification is semantically incorrect as he classifies more generalized concepts: place and accommodation as subclass of a concept Hotel. Figure 3 shows an error of this type while building semantic classification as class hierarchy in ontology. In other words rather doing specialization, he does generalization. This makes serious complications while interpretation of concepts and machine cannot reason about them effectively.

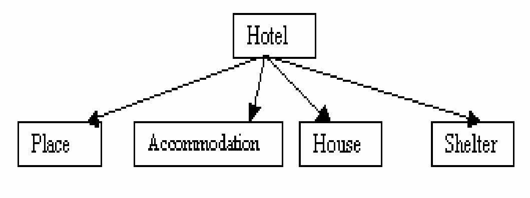

Figure 3. Weaker domain specified by subclasses Place and Accommodation

2.2. Domain breach specified by subclass error

These errors occur when ontologists specify subclass that possesses only few features of superclass and makes superclass domain breach by allowing more features that may not present in superclass or violate some features of superclass. Usually this type of error occurs when ontologists can not keenly observe all the features of concepts and classify them as a subclass of concept that only partially meet the properties of superclass. Ideally a subclass should possess all the features of its superclass and may add new features that should not conflict with its superclass features. For example consider a mathematical class hierarchy where ontologist classifies concepts: Even and Odd as subconcepts of a concept Natural Numbers. Furthermore he classifies prime and non-prime concepts as the subclasses of the concept Odd. This classification is semantically wrong, as not all prime numbers are Odd numbers, i.e. a digit 2 is a prime but not Odd. Figure 4 shows an error of this type while building semantic classification as class hierarchy in ontology.

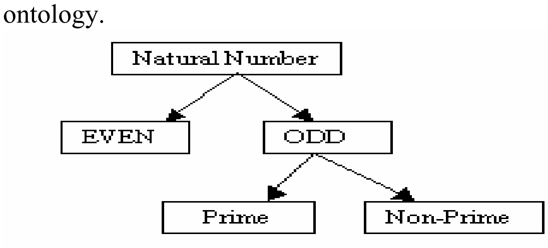

Figure 4. Domain breach specified by subclass Prime

2.3. Disjoint domain specified by subclass error

These errors occur when ontologists specify concepts as subclasses of a concept that occupies a disjoint domain. For example he classifies concepts Ship and Train as subclasses of Flight concept. None of the features of Ship and Train match with superclass concept Flight i.e. they belong to disjoint domains. Figure 5 shows an example of disjoint domain specified by subclasses Ship and Train.

These types of errors are not detected by the available evaluation systems like Racer, Fact and Pellet. They create the same problems as other inconsistency errors in ontology like inconsistency partition errors, so detection of these errors is as important as other inconsistency errors.

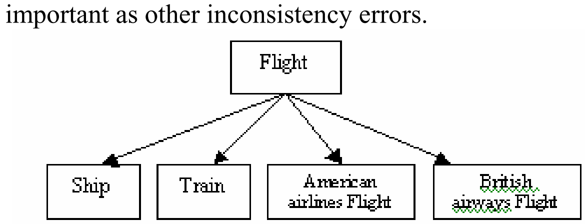

Figure 5. Disjoint domain specified by subclasses Ship and Train

3. Semantic inconsistency errors in Ontology Merging Process

Ontology Merging systems are the key factor for enabling interoperability across ontology-based heterogeneous systems and semantic web applications [13]. We developed DKP:OM, that uses many matching heuristics to produce initial mappings between source ontologies to aid user for consistent, coherent and complete merged ontology [11]. We experienced many semantic inconsistency errors, as different concepts of source ontologies map on one another at different places. All these kinds of errors occur when concepts in ontologies partially overlap with one another and merging system suggests user to merge concepts that are semantically close via equivalence and subsumption relations. Most common types of inter-ontology relationships in order to enable interoperability are Hyponym term (more specific than other) and Hypernym term (more general than other) [6]. Semantic inconsistency errors would increase if class hierarchies with in the source ontology do not well formed i.e. they have wrong semantic classification of concepts. As a result machine cannot be able to interpret concepts efficiently leading inconsistent merged ontology [10].

For well-formed class hierarchy, subclass should not belong to disjoint domain or weaken or breach the properties and features of any superclass concepts. The subclass may add new features and properties but they should specialize the features and properties of its super class. *WordNet* provides ontologist the list of synonyms, hyponyms and hyponyms. This helps a lot while semantic classification of concepts in the ontology and while validating mapping produced by the merging systems. To avoid semantic inconsistency, ontologists are suggested to make new subtypes and supertypes of concepts and relations in either ontology. Classification of semantic errors help ontologists to build ontologies that are free from semantic inconsistencies. The ontologists should not postpone the evaluation until the taxonomy is finished; the control mechanisms should be performed during the construction of the taxonomy or during finding similarities between concepts in ontology merging process. Identification of semantic inconsistency is highly significant because in its presence, ontologies lose their powerful mechanism for enhanced reasoning about concepts [2]. If taxonomies have semantic consistencies then machine cannot infer and reason about concepts effectively to achieve the goals of semantic web. Also we cannot be able to exploit combined use of independently developed ontologies and interoperability cannot be achieved. Semantic search engine is the major application of semantic web. The components of semantic search engine use ontology for their purpose like indexer semantically indexes the crawled pages by using ontology [7]. Semantic crawler component crawls the pages and finds semantic relevancy with domain by using ontology. Consider the situations where the concept of ontology itself has wrong classification of concepts. This will affect the results of semantic crawler and semantic indexer. As a result, we will be detracted from the goals of semantic web.

4. Conclusion

The main contribution of this paper is an extension in error's taxonomy. We have identified three types of semantic inconsistency errors. The main reason for these errors is that ontologists do not classify the concepts properly. They classify a concept as a subclass of another concept but these either have disjoint domains, or subclass concept breaches or weakens the superclass concepts. This classification helps to build well-formed class hierarchies that are free from semantic inconsistencies. This checklist provides ontologists to evaluate mapping produced by mapping/merging systems via equivalence and subsumption relations. We conclude that semantic inconsistencies should be detected to enable interoperability between ontologies and to achieve combined use of independent ontologies. In future work, we will further evaluate error's taxonomy and try to find some other type of errors that are usually encountered by ontologist for fulfillment of the sound Semantic Web vision.

5. References

[1] A. G´omez-P´erez et al, "Evaluation of Taxonomic Knowledge on Ontologies and Knowledge-Based Systems", International Workshop on Knowledge Acquisition, Modeling and Management (1999).

[2] G. Antoniou, F.V. Harmelen, "A Semantic Web Primer", MIT Press Cambridge 2004 ISBN 0-262-01210-3

[3] A. Gomez-Perez, M.F-Lopez, O.C-Garcia, "Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and the. Semantic Web", Published by Springer ISBN:1-85253-55j-3

[4] M.A. Qadir, W. Noshairwan, "Warnings for Disjoint Knowledge Omission in Ontologies" 2nd Intl. Conference on internet and Web Applications and Services (ICIW07). p. 45

[5] D. Nardi et al, "The Description Logic Handbook: Theory, Implementation, and Applications", ISBN: 9780521781763

[6] J. Baumeister and D.S. Smelly, "Owls–Design Anomalies in Ontologies", 18th Internaional Florida Artifiical Intelligence Research Society Conference (FLAIRS), pp 251-220, AAAI Press, 2005

[7] Ganesh, S. Jayaraj, M. Kalyan, V. SrinivasaMurthy Aghila, "Ontology-based Web crawler", Intl. conference of Information technology. ITCC 2004

[8] J. Brank et al, "A Survey of Ontology Evaluation Techniques", Published in multi-conference IS 2005, Ljubljana,Slovenia SIKDD 2005.

[9] W. Noshairwan, M. A. Qadir, M. Fahad, "Sufficient Knowledge Omission error and Redundant Disjoint Relation in Ontology", InProc. 5th Atlantic Web Intelligence Conference 2007June 25-27, 2007 - Fontainebleau, France

[10] C. Jelmini, and S. M-Maillet, "OWL-based reasoning with retractable inference", In RIAO Conference Proceedings, 2004.

[11] M. Fahad, M.A. Qadir, M.W. Noshairwan, N. Iftakhir, "DKP-OM: A Semantic Based Ontology Merger", Inproc 3rd Intl. Conf. i-semantics Austria 2007, publication J.UCS

[12] M. Klein, "Combining and relating ontologies: an analysis of problems and solution", In Gomez-Perez, A. et al., editors, Proceedings of Workshop on Ontologies and Information Sharing at the 17th IJCAI-2001, Seattle, USA.

[13] H. S. Pinto, A. Gomez-Perez, and J. P. Martins, ,"Some issues on ontology integration", In proceedings of the workshop on ontologies and Problem Solving Methods during IJCAI-99, Stockholm, Sweden, August 1999.