

Incompleteness Errors in Ontology

¹ Muhammad Abdul Qadir, ² Muhammad Fahad, ³ Syed Adnan Hussain Shah
Muhammad Ali Jinnah University, Islamabad, Pakistan
¹aqadir@jinnah.edu.pk, ²mhd.fahad@gmail.com, ³syedadnan@gmail.com

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 for fulfillment of the Semantic Web vision and classified them in error's taxonomy. We have found that some important errors are missing in the error's taxonomy. We have identified and defined two new incompleteness errors i.e. Functional Property Omission (FPO) for single valued property and Inverse-Functional Property Omission (IFPO) for a unique valued property. We have demonstrated the importance of such errors by giving different scenarios where appropriate. We have evaluated different ontologies and presented empirical results.

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. 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 [6].

Domain researchers have identified some errors and defined them in error's taxonomy for assistance in the ontology evaluation [1, 15]. This error's taxonomy becomes a guideline for ontology engineer to evaluate the ontology in perspective of such errors. If some

errors are not defined in error's taxonomy then we can say that the ontology engineer based on the error's taxonomy will not detect 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 [1, 3]. Redundancy means that same information is inferred from ontology more than once. Incompleteness means the concepts are not completely defined. Qadir et al [4] identified the scenarios where incomplete partition error such as disjoint knowledge omission leads toward catastrophic situations and proposed a system that detects disjoint knowledge omission and generates warnings for ontologists. Baumeister and Seipel [9] discussed the evaluation process of ontologies and design anomalies in ontology. They defined the detection method by using prolog and FN-query language. These anomalies help the ontologist to develop consistent ontology. Brank et al. [14] discussed the overall approaches for ontology evaluation and conclude that different approaches are useful in different application. The selection of evaluation approach depends on the application. This survey helps the ontologist to select appropriate evaluation method for its domain.

Previously we have extended Gomez's error taxonomy by identifying two new errors [15]. 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 but in actual they are already disjoint as their parents are disjoint.

In this paper we have shown our current findings about the evaluation of error taxonomy. We identified that some important incompleteness errors are missing in error's taxonomy i.e. Functional Property Omission (FPO) for single valued property and Inverse Functional Property Omission (IFPO) for a unique valued property. If these errors are not detected then it

is impossible to specify property characteristics that provide a powerful mechanism for enhanced reasoning about a property. Moreover their presence in ontology creates several inconsistencies and weakens inferring mechanisms. 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 presents classification of errors and our contribution to error's taxonomy; section 3 presents the empirical results. Section 4 concludes the paper and gives insight on future work.

2. Extension in Error's Taxonomy

In OWL, it is possible to specify property characteristics that provide a powerful mechanism for enhanced reasoning about a property [2,7]. OWL functional and inverse-functional properties indicate how many times a property can be used for a given subject or object. Sometimes ontologists do not give significance to these properties and do not declare datatype or object properties as functional or inverse-functional. This results machine not to reason about a property effectively leading to serious complications.

2.1 Functional Property Omission (FPO) for single valued property

A functional property is one that has just one value for any particular subject. A formal definition of functional property [17] is: If a property, P , is tagged as *functional* then for all x , y , and z : $P(x,y)$ and $P(x,z)$ implies $y = z$. Both object properties and datatype properties can be declared as "functional". An example of object property is the *hasBirthday* relation between a person and his or her birthday. Everyone has just one birthday, so for any given subject (person), there can be just one object (birthday). Likewise datatype properties can be declared as "functional" that can have only one value y for each instance x , i.e. there cannot be two distinct values y_1 and y_2 such that the pairs (x,y_1) and (x,y_2) are both instances of this property. An example is the *Rank_No* datatype property that belongs to the student concept. Every student has just one *Rank_No* in the department, so for any given subject (student), there can be just one object (*Rank_No*).

Sometimes ontologists unintentionally do not mark a single valued property as functional. As a result property accepts multiple values and thus ontologists have created inconsistency within the ontology due to incomplete specification of concepts. The main reason

for this is that OWL by default supports multi-values for datatype property and object property, and ontologists did not paid attention to this fact. We consider such lack of information as an error, because such ignorance leads towards serious complications. Section 2.3 presents how FPO detracted to achieve the desired goals and create inconsistencies.

2.2 Inverse-Functional Property Omission (IFPO) for a unique valued property

If a property is declared to be inverse-functional, then the object of a property statement uniquely determines the subject (some individual). Inverse-functional properties resemble the notion of a unique key in databases [17]. A formal definition of inverse-functional property is: If a property, P , is tagged as *InverseFunctional* then for all x , y and z : $P(y,x)$ and $P(z,x)$ implies $y = z$. This means that if we state P as an owl *InverseFunctionalProperty*, then this asserts that a value x can only be the value of P for a single instance, i.e. there cannot be two distinct instances y and z such that both pairs (y, x) and (z, x) are instances of P . In OWL Full datatype properties are a subclass of object properties, and an inverse-functional property can be defined for datatype properties as well. But in OWL DL object properties and datatype properties are disjoint, so an inverse-functional property cannot be defined for datatype properties. An example of a property ISBN is inverse-functional for a object book, as there is only one subject for a particular inverse functional property. We mean that for any ISBN, there is only one book that has that ISBN.

Sometimes ontologists unintentionally do not mark a property as inverse functional. As a result, they have created inconsistency within the ontology due to incomplete specification of concepts. We consider such lack of information as an error, because such ignorance leads machine not to infer and reason about concepts uniquely.

2.3 Consequence of Functional Property Omission and Inverse-Functional Property Omission

We describe the importance of these errors by defining some scenarios, where proper results or objectives are not achieved due to presences of these incompleteness errors in ontology.

- **Inconsistency in Ontology:** These incompleteness errors lead towards inconsistency within the ontology. Omission of functional property for single valued datatype property or

object property makes the wrong interpretation about the modeled domain concepts. Consider an ontology where ontologists unintentionally do not make *hasBirthday* property as functional property. Such lack of information creates inconsistency as it allows someone to have many birthdays. Likewise inverse-functional property omission for ISBN would create inconsistency by letting many ISBN numbers to identify one book.

- **Inferring Concepts of Ontology:** Functional and inverse-functional properties can be used by applications to infer things like two entities denote the same thing. For example consider the scenario where the *has_National_ID* property is inversefunctional then the facts (Person:a *has_National_ID* 314-3943-6) and (Person:b *has_National_ID* 314-3943-6), let the application to conclude that Person:a and Person:b denote the very same person. Inferring concepts by such information helps in the process of finding mapping between two ontologies on the basis of instances. Consider another scenario [16] where the *hasMother* property is functional then the facts (Philippe *hasMother* XYZ) and (Philippe *hasMother* ABC) imply that (ABC owl:sameIndividualAs XYZ). When ontologists ignore to declare a property as functional or inverse-functional where it requires then applications do not able to infer things like whether two entities denote the same thing.
- **Mapping/Merging of Ontology:** The description of concepts is the most important aspect during the merging process of ontologies [18]. The merging system finds the similarities between concepts on the basis of concept's description and other information. The concept's description plays an important role to find correspondences among concepts. On the basis of found correspondences, ontologies are merged together to able interoperability among heterogeneous ontologies [19]. If concepts have lack of functional property and inverse functional property where they require then basically we did not get the advantages of OWL concept's description richness in merging process. As a result the process needs human intervention for completion or produced results may be less accurate.
- **Semantic search engine:** Semantic search engine is the fundamental application of semantic web. The component of semantic search engine use ontology for their purpose like indexer

semantically indexes the crawled pages by using ontology [12]. Semantic crawler component crawls the pages and finds semantic relevancy with domain by using ontology. Consider the situations where the concepts of ontology itself have such incompleteness errors. This will affect the results of semantic crawler and semantic indexer. As a result, we will be detracted from the goals of semantic web.

The above scenarios describe the significant importance of the errors and show that if we do not consider the error then we have to face some critical problems to achieve the objectives. These types of errors are not defined in the error's taxonomy and the available evaluation systems like Racer [8], Fact [10] and Pellet [11] do not detect them. It will create same problems as other incompleteness errors in the ontology, so detection of these errors is as important as detection of other incompleteness errors. But the detection of such errors is quite difficult. One of the main problems due to which evaluation systems do not detect such incompleteness is that they do not interpreter semantics. It is human who knows that for any given subject (person), there can be just one object (birthday) or for any object (book), there is only one unique subject (ISBN). One of the possible ways to detect such incompleteness errors is to evaluate ontology on test data (valid and invalid both) that is generated according to tester's domain knowledge, experience with similar concepts and information about soft spots of ontology. As this approach is totally dependant on best engineering judgment, hence detection of such errors is very dependant on the abilities of the tester.

3. Empirical Results

We have evaluated some known ontologies [13] to prove that ontologists usually commit functional property omission error for single valued property and inverse-functional property omission error for a unique valued property. After evaluation we have found that most of the ontologies have such (FPO and IFPO) incompleteness errors. We call concepts having incompleteness errors as warned concepts. The summary of evaluation results is shown in Table 1. We have evaluated ontologies by testing wrong input instances against the concepts of these ontologies. Properties missing functional and inverse-functional declaration accept multiple values that can be used for a given concept, generating inconsistencies within the ontology.

Table 1. Empirical results

Ontology	Number of Warned Concepts	Total Number of Concepts
Pizza	12	96
Camera	3	12
Generations	4	18
People-Pets	11	70
Travel	7	35

4. Conclusion

The main contribution of this paper is an extension in error's taxonomy. We have identified two new types of incompleteness errors i.e. Functional Property Omission (FPO) for single valued property and Inverse Functional Property Omission (IFPO) for a unique valued property. We have also described the importance of detection of these errors by explaining different scenarios. We evaluated different ontologies and found that the FPO and IFPO errors are present in them. In future work, we will give the detection method to find these incompleteness errors to enable the fulfillment of the Semantic Web vision and further evaluate error's taxonomy and try to find some other type of errors that are usually encountered by ontologist.

5. References

- [1] A. Gómez-Pérez 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. Fernandez-Lopez, A. Gomez-Perez, O. Corcho-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, M.W. Noshairwan "Warnings for Disjoint Knowledge Omission in Ontologies", Second International Conference of 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] DAML: Available. <http://www.daml.org/> (Aug 2007)
- [7] Web Ontology Language Overview: Available. <http://www.w3.org/TR/owl-features/> (current Aug 2007)
- [8] V. Haarslev, and R. Moller, "RACER System Description", In Proceeding of the International Joint Conference on Automated Reasoning ,IJCAR'2001, pp 701-705, LNCS, Springer-Verlag, 2001.
- [9] J. Baumeister, "Smelly Owls-Design Anomalies in Ontologies", 18th Intl. Florida Artificial Intelligence Research Society Conference, pp 251-220, AAAI Press, 2005
- [10] I. Horrocks, "The FaCT System." International conference. on Analytic Tableaux and Related Methods, pp 307-312, vol 1397, Springer-Verlag, 1998
- [11] Pellet, "An OWL DL Reasoner", Available. www.pellet.owdl.com/ (Current Aug 2007)
- [12] Ganesh, S. Jayaraj, M. Kalyan, V. Srinivasa Murthy Aghila., "Ontology-based Web crawler", International conference of Information technology. ITCC 2004
- [13] Protégé Ontologies Library. Available: <http://protege.cim3.net/cgi-bin/wiki.pl?ProtegeOntologiesLibrary#nid81S> (current August 2007).
- [14] J. Brank, et al., "A Survey of Ontology Evaluation Techniques", Published in multi-conference IS 2005, Ljubljana, Slovenia SIKDD 2005.
- [15] W. Noshairwan, M. A. Qadir, M. Fahad, "Sufficient Knowledge Omission error and Redundant Disjoint Relation in Ontology". In Proc. 5th Atlantic Web Intelligence Conference 2007 June 25-27, 2007 - Fontainebleau, France
- [16] C. Jelmini, and S. M-Maillet, "OWL-based reasoning with retractable inference", In RIAO Conference Proceedings, 2004.
- [17] Ontology Definition Metamodel, Sixth Revised Submission to OMG/RDF ad/2005-04-13
- [18] A. H. Doan, "Learning to map between structured representations of data", PhD thesis, University of Washington, Seattle (WA US), 2002.
- [19] M. Fahad, M.A. Qadir, M.W. Noshairwan, N. Iftakhir, "DKP-OM: A Semantic Based Ontology Merger", Inproc. 3rd Intl. Conference i-semantics 2007, publication at J.UCS