Metrics for Evaluating the Semantic Implications of Changes in Evolving Ontologies

Chuming Chen and Manton M. Matthews
Department of Computer Science and Engineering
University of South Carolina
Columbia, SC 29208, USA

Abstract *We propose a set of metrics for measuring the semantic implications of changes during ontology evolution. Our metrics focus on the changes of classes and associated axioms or annotations in an evolving OWL ontology. We identify three categories of changes and present class, class definition, class subsumption hierarchy, and logical model oriented metrics for quantifying these changes. We implement the proposed metrics using OWL API and Pellet OWL-DL reasoner. We evaluate our metrics by applying them to the evolution of the Gene Ontology.*

Keywords: Change Evaluation Metrics, OWL, Semantic Implications, Ontology Evolution, Inference

1 Introduction

OWL [5] is a family of Description Logics (DL) [2] based ontology languages designed to enable automated machine reasoning on the Semantic Web. OWL-DL is a rich ontology language that supports high expressiveness and decidable reasoning. An OWL-DL ontology corresponds to a $\mathcal{SHOIN}(\mathcal{D})$ [6] knowledge base that consists of a set of axioms and annotations, which describe/define a set of classes, properties and individuals. An individual is a single object in a domain. A class is a group of objects in a domain. A property represents a binary relationship between two classes.

The dynamic nature of the application domains and the imperfect knowledge acquisition process suggest that ontology developers unavoidably need to revise an ontology. The objectives of this work are to evaluate the dynamics of ontologies in terms of the semantic implications of changes and to help ontology developers understand the scale and consequences of the changes they have made. We propose a set of semantics-aware metrics for the changes in the subsequent versions of an evolving ontology. We implement the proposed metrics using OWL API [4] and Pellet [8] OWL-DL reasoner.

The rest of the paper is structured as follows: Section 2 discusses the semantic implications of changes in evolving OWL ontologies. Section 3 introduces our ontology change evaluation metrics. Section 4 evaluates our metrics using the evolution of the Gene Ontology [3] as an example. Section 5 reviews related work. We conclude the paper and outline the future work in Section 6.

2 Semantic Implications of Changes

An ontology is a shared conceptualization of a domain in a formal language. It uses a set of axioms to explicitly make statements that say what is true in the domain being modeled. In an evolving ontology, we can identify two levels of changes based on this model: a) *Changes of Entities*, which are represented by adding or removing classes in the ontology. This type of change indicates which domain entities are being modeled by the ontology, and captures the high-level changes in the domain. b) *Changes of Entity Semantics*, which are represented by adding or removing axioms which describe certain aspects of a class in the ontology. This type of change reflects how the domain entities are being modeled by the ontology, and captures the low-level changes in the domain.

An OWL ontology comprises a set of axioms and annotations. The axioms encompass the semantics of the ontology and include conventional DL axioms and assertions. The annotations have no logical meanings and are only included to provide human-friendly documentation for the ontology [4]. It is important that we distinguish the logical changes that affect the logical theory underlying the ontology from changes of annotations that have no semantic implications: a) *Logical Changes*, which are represented by adding or removing logical axioms in the ontology. This type of change has logical consequence to the ontology. In other words, they may change the classification hierarchy of named classes and lead to different query results. b) *Non-logical Changes*, which are represented by adding or removing annotations in the ontology. This type of change does not have any logical consequence.

In an OWL ontology, in addition to the explicitly stated axioms, we can use a reasoner to infer additional logical axioms that are entailed by the ontology. Therefore, we need to understand whether changes are made to the explicit or implicit model of the ontology: a) *Asserted Changes*, which are represented by adding or removing the explicitly stated axioms or annotations in the ontology. This type of change can be computed by comparing the structures of two ontologies. b) *Inferred Changes*, which are represented by adding or removing the inferred axioms that are entailed by the ontology. This type of change can only be computed as the result of a reasoning process. Note that non-logical changes can not generate inferred changes.

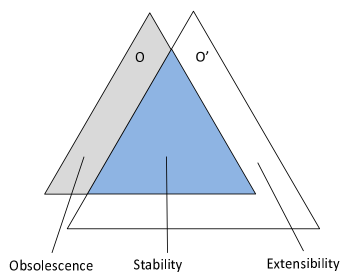

Figure 1: Ontology Change Effects Model

In this paper, we measure the ontology changes with respect to the following ramification of the evolving ontology model (Figure 1): a) *Stability* shows how much of the ontology remain unchanged when it is turning into a new ontology. b) *Extensibility* shows how much of the ontology is new after it has been changed. c) *Obsolescence* shows how much of the ontology has become obsolete.

3 Change Evaluation Metrics

Our metrics focus on measuring the ontology changes and their semantic implications from the following perspectives:

Class Oriented Metrics that examine the classes added to or removed from an ontology. These metrics indicate the high-level changes to the ontology model.

Class Definition Oriented Metrics that examine the changes in the class definition, which is a set of logical axioms (both asserted and inferred) and non-logical annotations reference this particular class. These metrics reveal the low-level fine-grained changes to the ontology model.

Class Subsumption Hierarchy Oriented Metrics that examine the changes of maximum depth, average depth, fan-in and fan-out of named classes in the hierarchy (Figure 2). These metrics show the changes of completeness and breadth of an evolving ontology. We also measure the changes of depths of the stable classes in the class subsumption hierarchies. This measure indicates whether a named class has been generalized, specialized or no change at all with respect to the entire hierarchy (Figure 3).

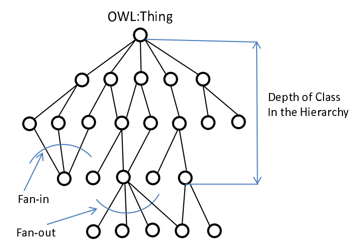

Figure 2: Depth, Fan-in and Fan-out of Named Classes in the Class Subsumption Hierarchy

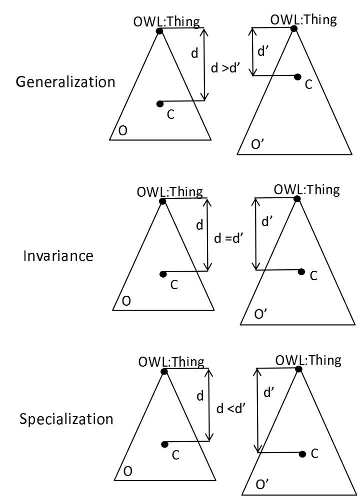

Figure 3: Changes of the Stable Classes in the Class Subsumption Hierarchy

Logical Model Oriented Metrics that examine the proportion of logical axioms and non-logical annotations changed, which characterizes the impact of changes on the semantic model of the ontology.

We now present a set of change evaluation metrics characterizing the changes from the above perspectives. To make the metrics comparable, we also define the normalized versions of the change metrics whenever appropriate. Note that these metrics are defined based on the ontologies that have incorporated the inferred axioms. i.e. they are the logical closure of the original ontologies.

3.1 Class Oriented Metrics

These metrics measure the stability, extensibility and obsolescence of named classes between two versions of an ontology O and O'.

Metric 1 Class Stability (CS)

$$\begin{aligned} CS(O \rightarrow O') &= \|\{c \mid c\, is\, a\, named\, class,\ c \in O\, and\, c \in O'\}\| \\ NoC(O) &= \|\{c \mid c\, is\, a\, named\, class,\ c \in O\}\| \\ NCS(O \rightarrow O') &= CS(O \rightarrow O')/NoC(O) \end{aligned}$$

Metric 2 Class Extensibility (CE)

$$\begin{aligned} CE(O \rightarrow O') &= \|\{c \mid c\, is\, a\, named\, class,\ c \notin O,\ but\, c \in O'\}\| \\ NoC(O') &= \|\{c \mid c\, is\, a\, named\, class,\ c \in O'\}\| \\ NCE(O \rightarrow O') &= CE(O \rightarrow O')/NoC(O') \end{aligned}$$

Metric 3 Class Obsolescence (CO)

$$\begin{aligned} CO(O \rightarrow O') &= \|\{c \mid c\, is\, a\, named\, class,\ c \in O,\ but\, c \notin O'\}\| \\ NoC(O) &= \|\{ c \mid c\, is\, a\, named\, class,\ c \in O\}\| \\ NCO(O \rightarrow O') &= CO(O \rightarrow O')/NoC(O) \end{aligned}$$

3.2 Class Definition Oriented Metrics

These metrics measure the stability, extensibility and obsolescence of the definitions for the stable classes within an ontology O' with respect to its previous version O. The definition of a named class c in the ontology O, $DefA(c, O)$, consists of a set of axioms (asserted and inferred) and annotations that describe/define the class c. In other words, any axiom or annotation references the class c.

Metric 4 Class Definition Stability (CDS)

$$\begin{aligned} CDS(O \rightarrow O') &= \|\{ c \mid c\, is\, a\, named\, class,\ c \in O \\ &\quad and\, c \in O', \forall\, a,\ a\, is\, an\, axiom \\ &\quad or\, annotation,\ if\, a \in DefA(c, O), \\ &\quad then\, a \in DefA(c, O')\}\| \\ NCDS(O \rightarrow O') &= CDS(O \rightarrow O')/CS(O \rightarrow O') \end{aligned}$$

$CS(O \rightarrow O')$ is defined in the metric 1.

Metric 5 Class Definition Extensibility (CDE)

$$\begin{aligned} CDE(O \rightarrow O') &= \|\{ c \mid c\, is\, a\, named\, class,\ c \in O \\ &\quad and\, c \in O',\ \exists\, a,\ a\, is\, an\, axiom \\ &\quad or\, annotation,\ a \notin DefA(c, O), \\ &\quad but\, a \in DefA(c, O')\}\| \\ NCDE(O \rightarrow O') &= CDE(O \rightarrow O')/CS(O \rightarrow O') \end{aligned}$$

$CS(O \rightarrow O')$ is defined in the metric 1.

Metric 6 Class Definition Obsolescence (CDO)

$$\begin{aligned} CDO(O \rightarrow O') &= \|\{ c \mid c\, is\, a\, named\, class,\ c \in O \\ &\quad and\, c \in O',\ \exists\, a,\ a\, is\, an\, axiom \\ &\quad or\, annotation,\ a \in DefA(c, O), \\ &\quad but\, a \notin DefA(c, O')\}\| \\ NCDO(O \rightarrow O') &= CDO(O \rightarrow O')/CS(O \rightarrow O') \end{aligned}$$

$CS(O \rightarrow O')$ is defined in the metric 1.

3.3 Class Subsumption Hierarchy Oriented Metrics

We can use a reasoner to compute the partial ordering or subsumption hierarchy of named classes in an OWL ontology. It is also called classification. The structure of class subsumption hierarchy represents the completeness and breadth of the domain knowledge and can be characterized by the notions of depth, fan-in and fan-out (Figure 2). The hierarchy usually starts from the most general class or class without any superclasses (the concept $\top$ in DL or OWL:Thing in OWL). Given a class subsumption hierarchy of an ontology O, the depth of a named class c in the hierarchy, $Depth(c, O)$, is defined as the shortest path length from the class OWL:Thing to this class. The fan-in of a named class c in the hierarchy, Fan-in(c, O), is defined as the total number of its superClasses. The fan-out of a named class c in the hierarchy, Fan-out(c, O), is defined as the total number of its subClasses.

Metric 7 Maximum Depth (MD)

$$MD(O) = max(Depth(c, O), \forall c \in O)$$

The MD indicates how deep a named class can be in the hierarchy. In other words, how many levels of generalities we have for the domain knowledge.

Metric 8 Average Depth (AD)

$$AD(O) = \sum (Depth(c, O), \forall c \in O)/NoC(O)$$

where $NoC(O)$ denotes the total number of classes as defined in the metric 1.

The AD shows on average how deep a named class can be in the hierarchy.

Metric 9 Maximum Fan-in (MFI)

$$MFI(O) = max(Fan\text{-}in(c, O), \forall c \in O)$$

The MFI shows the maximum number of multiple inheritance a named class can have in the hierarchy.

Metric 10 Average Fan-in (AFI)

$$AFI(O) = \sum(Fan\text{-}in(c, O), \forall c \in O)/NoC(O)$$

where $NoC(O)$ denotes the total number of classes as defined in the metric 1.

The AFI shows on average how many multiple inheritances a named class can have in the hierarchy.

Metric 11 Maximum Fan-out (MFO)

$$MFO(O) = max(Fan\text{-}out(c, O), c \in O)$$

The MFO shows how wide a branch can be or the maximum number of subclasses a named class can have in the hierarchy.

Metric 12 Average Fan-out (AFO)

$$AFO(O) = \sum(Fan\text{-}out(c, O), \forall c \in O)/NoC(O)$$

where $NoC(O)$ denotes the total number of classes as defined in the metric 1.

The AFO shows on average how wide a branch can be in the hierarchy.

The change of depth of a named stable class in the hierarchy indicates whether it has been generalized or specialized with respect to the entire hierarchy as shown in Figure 3. Generalization of a named class means the expansion of its extension, i.e. more objects in the domain can be classified as instances of that class. Specialization of a named class means the contraction of its extension, i.e. less objects in the domain can be classified as instances of this named class. By measuring this property, we can quantify the semantic consequences of changes to the ontology.

Metric 13 Generalized Classes (GC)

$$\begin{aligned} GC(O \to O') &= \|\{c \mid c\ is\ a\ named\ class,\ c \in O\ and\ c \\ &\quad \in O',\ Depth(c, O) > Depth(c, O')\}\| \\ NGC(O \to O') &= GC(O \to O')/CS(O \to O') \end{aligned}$$

$CS(O \to O')$ is defined in the metric 1.

Metric 14 Specialized Classes (SC)

$$\begin{aligned} SC(O \to O') &= \|\{c \mid c\ is\ a\ named\ class,\ c \in O\ and\ c \\ &\quad \in O',\ Depth(c, O) < Depth(c, O')\}\| \\ NSC(O \to O') &= SC(O \to O')/CS(O \to O') \end{aligned}$$

$CS(O \to O')$ is defined in the metric 1.

Metric 15 Invariant Classes (IC)

$$\begin{aligned} IC(O \to O') &= \|\{c \mid c\ is\ a\ named\ class,\ c \in O\ and\ c \\ &\quad \in O',\ Depth(c, O) = Depth(c, O')\}\| \\ NIC(O \to O') &= IC(O \to O')/CS(O \to O') \end{aligned}$$

$CS(O \to O')$ is defined in the metric 1.

For the changes of locations of classes in the class subsumption hierarchy, we can also measure whether changes occurred to the classes that are in the leaf positions of a hierarchy. We call them leaf classes (classes without any subclasses). Adding or removing those classes have the minimum semantic impact on the entire classification hierarchy.

Metric 16 Leaf Class Stability (LCS)

$$\begin{aligned} LCS(O \to O') &= \|\{c \mid c\ is\ a\ named\ leaf\ class,\ c \in O \\ &\quad and\ c \in O'\}\| \\ NLCS(O \to O') &= LCS(O \to O')/CS(O \to O') \end{aligned}$$

$CS(O \to O')$ is defined in the metric 1.

Metric 17 Leaf Class Extensibility (LCE)

$$\begin{aligned} LCE(O \to O') &= \|\{c \mid c\ is\ a\ named\ leaf\ class,\ c \notin O, \\ &\quad but\ c \in O'\}\| \\ NLCE(O \to O') &= LCE(O \to O')/CE(O \to O') \end{aligned}$$

$CE(O \to O')$ is defined in the metric 2.

Metric 18 Leaf Class Obsolescence (LCO)

$$\begin{aligned} LCO(O \to O') &= \|\{c \mid c\ is\ a\ named\ leaf\ class,\ c \in O, \\ &\quad but\ c \notin O'\}\| \\ NLCO(O \to O') &= LCO(O \to O')/CO(O \to O') \end{aligned}$$

$CO(O \to O')$ is defined in the metric 3.

3.4 Logical Model Oriented Metrics

These metrics measure the stability, extensibility and obsolescence of logical axioms or non-logical annotations between two versions of an ontology O and O'.

Metric 19 Logical Axiom Stability (LAS)

$$\begin{aligned} LAS(O \to O') &= \|\{a \mid a\ is\ a\ logical\ axiom,\ a \in O \\ &\quad and\ a \in O'\}\| \\ NoLAA(O) &= \|\{a \mid a\ is\ an\ axiom\ or\ annotation, \\ &\quad a \in O\}\| \\ NLAS(O \to O') &= LAS(O \to O')/NoLAA(O) \end{aligned}$$

Metric 20 Logical Axiom Extensibility (LAE)

$$\begin{aligned} LAE(O \to O') &= \|\{a \mid a\ is\ a\ logical\ axiom,\ a \notin O, \\ &\quad but\ a \in O'\}\| \\ NoLAA(O') &= \|\{a \mid a\ is\ an\ axiom\ or\ annotation, \\ &\quad a \in O'\}\| \\ NLAE(O \to O') &= LAE(O \to O')/NoLAA(O') \end{aligned}$$

Metric 21 Logical Axiom Obsolescence (LAO)

$$LAO(O \rightarrow O') = \|\{a \mid a\ is\ a\ logical\ axiom,\ a \in O,\ but\ a \notin O'\}\|$$
$$NLAO(O \rightarrow O') = LAO(O \rightarrow O')/NoLAA(O)$$

$NoLAA(O)$ is defined in the metric 19.

Metric 22 Annotation Stability (AS)

$$AS(O \rightarrow O') = \|\{a \mid a\ is\ an\ annotation,\ a \in O\ and\ a \in O'\}\|$$
$$NAS(O \rightarrow O') = AS(O \rightarrow O')/NoLAA(O)$$

$NoLAA(O)$ is defined in the metric 19.

Metric 23 Annotation Extensibility (AE)

$$AE(O \rightarrow O') = \|\{a \mid a\ is\ an\ annotation,\ a \notin O,\ but\ a \in O'\}\|$$
$$NAE(O \rightarrow O') = AE(O \rightarrow O')/NoLAA(O')$$

$NoLAA(O')$ is defined in the metric 20.

Metric 24 Annotation Obsolescence (AO)

$$AO(O \rightarrow O') = \|\{a \mid a\ is\ an\ annotation,\ a \in O,\ but\ a \notin O'\}\|$$
$$NAO(O \rightarrow O') = AO(O \rightarrow O')/NoLAA(O)$$

$NoLAA(O)$ is defined in the metric 19.

4 Evaluation

The Gene Ontology (GO) is a well-maintained and highly used biomedical ontology. It has three organizing sub-ontologies: biological process, cellular component and molecular function. The Gene Ontologies in the OBO-edit[1] format from Nov. 2006 to Jan. 2008 were downloaded from the GO archive ftp site[2]. A GO to OWL converter[3] was used to transform a GO file in the OBO-edit format into the OWL-DL format. We developed a tool called OntoCM on top of the OWL API [4] and Pellet OWL-DL reasoner [8] to measure the aforementioned change metrics for each of the three organizing sub-ontologies of the evolving Gene Ontology. Due to limited space, we only present the results for the biological process sub-ontology. The results for the other two are similar.

Class Oriented Changes. Knowledge acquisition is an accumulating process. This is clearly shown in Figure 4A. The total number of stable classes in an evolving ontology is steadily growing. As more research projects have been explored, our understanding of biological domain is also expanded. Although, we inevitably make changes to the

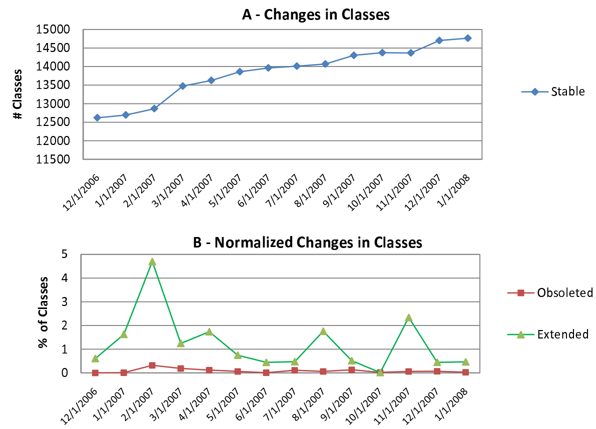

Figure 4: Changes of Classes

ontology due to incomplete and imperfect knowledge acquisition practice. However, those changes are not so significant, e.g. on average, less than 0.1% of classes were removed from the ontology, and about 1.2% classes were added to the ontology (Figure 4B).

Class Definition Oriented Changes. These metrics only consider the changes of definitions for the stable classes. In other words, we don't count the changes induced by the removed or added classes. We can see that the total number of classes with stable definitions is also steadily increasing (Figure 5A and Figure 4A), which indicates that our knowledge about the biological entities are stable and further research does not significantly invalidate our previous knowledge. On average, 3.4% of stable classes had their definitional axioms or annotations removed from the ontology, 0.8% stable classes had their definitional axioms or annotations extended (Figure 5B).

There are some irregular changes that happened during Jan. 2007 and Feb. 2007. The Gene Ontology CVS log and Monthly Report indicate that there were significant changes to the annotations of classes. This can be confirmed from Figure 9 and Figure 10. A large chunk of work was done offline and incorporated back all at once.

Class Subsumption Hierarchy Oriented Changes. In spite of the abnormal activity during Jan. 2007 and Feb. 2007, we can still observe that the average depth (Figure 6A), the average fan-in (Figure 6B), and the average fan-out (Figure 6C) of class subsumption hierarchy remain almost the same. These observations indicate that the coverage of the domain knowledge by the ontology is quite stable. Changes of each iteration only induce internal reorganization of classification of our knowledge base. This can be confirmed from the changes of locations for stable classes in the hierarchy as shown in Figure 7A. Only very few stable classes got moved up or down in the hierarchy. On average, about 56% of stable classes are leaf classes, 57% and 62% of obsolete and extended classes are leaf classes respectively (Figure 8B).

[1] http://www.geneontology.org/GO.format.obo-1_2.shtml
[2] ftp://ftp.geneontology.org/pub/go/ontology-archive/
[3] http://www.gong.manchester.ac.uk/

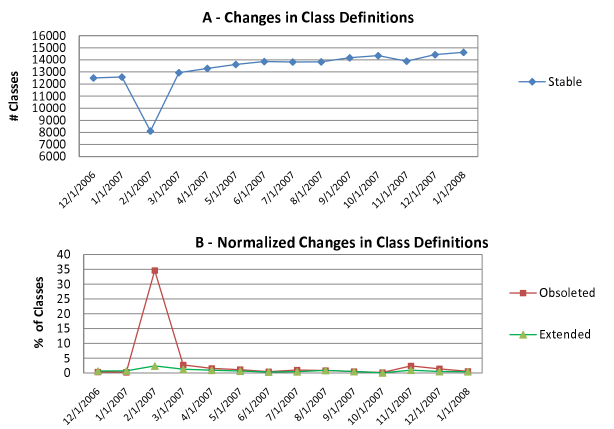

Figure 5: Changes of Class Definitions

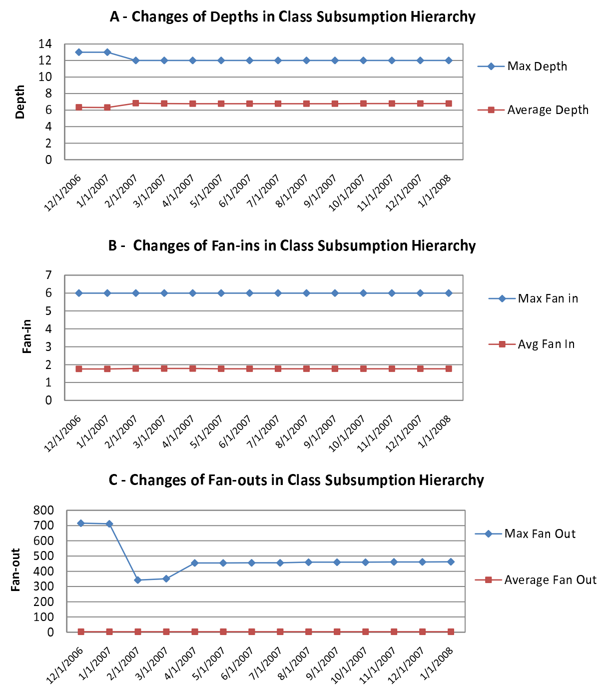

Figure 6: Changes of Depths, Fan-ins and Fan-outs in the Class Subsumption Hierarchy

Logical Model Oriented Changes. The total number of logical axioms and annotations is also gradually increasing and stable (Figure 9A, 10A). Changes of annotations almost double the contribution to the overall changes (Figure 9B, 10B). This suggests that most of changes were either adding or correcting annotations in the ontology, no significant changes to the logical model of the ontology.

5 Related Work

Vrandecic and Sure [9] raised the issue of incorporating ontology semantics when creating ontology metrics and proposed a principal approach to define ontology metrics

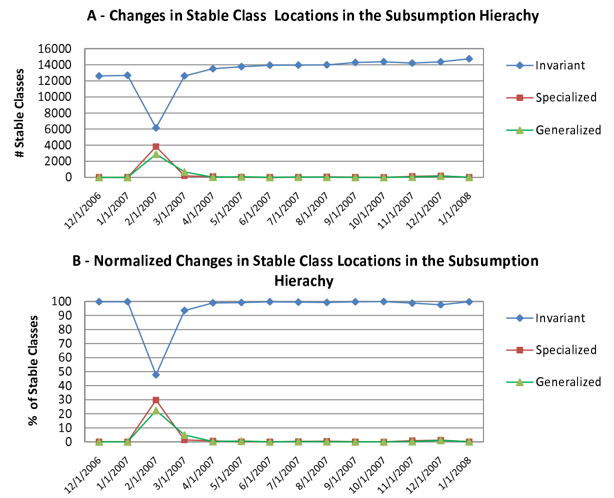

Figure 7: Changes of Stable Class Locations in the Class Subsumption Hierarchy

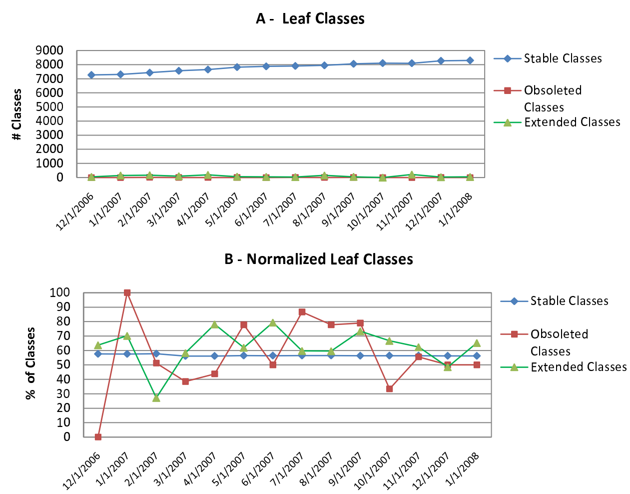

Figure 8: Changes of Leaf Classes in the Class Subsumption Hierarchy

based on the notions of "normalization" and "stable metrics". Our metrics are defined on the logical closure of the ontology model, therefore they satisfy these requirements.

Avery and Yearwood [1] identified the changes in the evolving ontology as two types: 1) Translocation Changes, which record which entities are modeled; 2) Transformation Changes, which record how an entity is modeled. Their work motivated our class oriented and class definition oriented change metrics.

The metrics reported here is comparable to the work by [7]. Their framework uses a temporal logic based multi-version ontology reasoning system to examine and analyze the ontology changes and their effects. The most notable difference is that we compute the semantic closure of an ontology and measure the changes of the class subsumption hierarchy. We also distinguish the changes

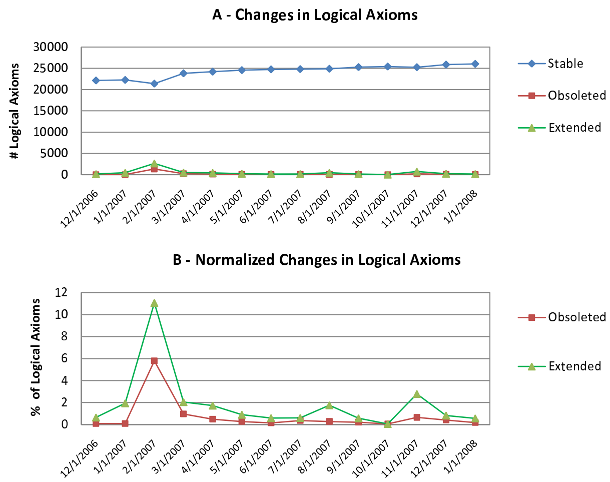

Figure 9: Changes of Logical Axioms

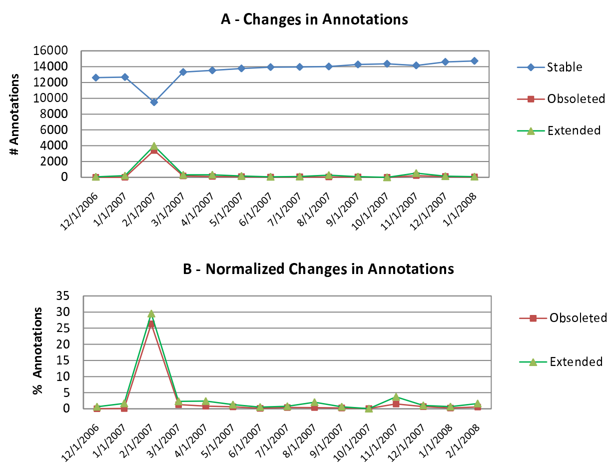

Figure 10: Changes of Annotations

that have logical consequences from those that don't. Our metrics measure the semantic implications of changes from class oriented, class definition oriented, class subsumption hierarchy oriented, and logical model oriented perspectives.

6 Conclusions and Future Work

We identified three categories of changes and presented a set of class, class definition, class subsumption hierarchy, and logical model oriented change metrics for quantifying these changes. We implemented the proposed metrics and performed an empirical study.

Our study of the Gene Ontology evolution revealed that changes happened very often, but on average, less than 0.1% of classes were removed from the ontology, and about 1.2% classes were added to the ontology. On average, 3.4% of stable classes had their definitional axioms or annotations removed from the ontology, 0.8% stable classes had their definitional axioms or annotations extended. On average, about 56% of stable classes are leaf classes, 57% and 62% of obsolete and extended classes are leaf classes respectively. Majority of changes are related to annotations of the ontology, which have no logical consequence to the ontology model.

The primary future work is to do further evaluation using real-world evolving OWL ontologies, especially ontologies that involve more dynamic changes in the properties and individuals of the ontology.

References

[1] John Avery and John Yearwood. "Supporting Evolving Ontologies By Capturing the Semantics of Change". In The Tenth Australian World Wide Web Conference, 2004.

[2] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi and Peter F. Patel-Schneider. "The Description Logic Handbook: Theory, Implementation and Application". Cambridge University Press, 2003.

[3] Gene Ontology Consortium. "The Gene Ontology (GO) Database and Informatics Resource". Nucleic Acids Research, 32 (Database-Issue): 258-261, 2004.

[4] Matthew Horridge, Sean Bechhofer and Olaf Noppens. "Igniting the OWL 1.1 Touch Paper: The OWL API". In 3rd OWL Experienced and Directions Workshop, 2007.

[5] Ian Horrocks, Peter F. Patel-Schneider and Frank van Harmelen. "From $\mathcal{SHIQ}$ and RDF to OWL: The Making of a Web Ontology Language". Journal of Web Semantics, 1(1): 7-26, 2003.

[6] Ian Horrocks and Ulrike Sattler. "A Tableaux Decision Procedure for $\mathcal{SHOIQ}$". In Proceedings of the 19th International Joint Conference on Artificial Intelligence(IJCAI), p. 448-453, 2005.

[7] Zhisheng Huang, Stefan Schlobach, Frank van Harmelen, Michel Klein, Nuria Casellas and Pompeu Casanovas. "Reasoning with Multi-version Ontologies: Evaluation, SEKT deliverable d3.5.2". Vrije Universiteit Amsterdam, 2006.

[8] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur and Yarden Katz. "Pellet: A Practical OWL-DL Reasoner". Journal of Web Semantics, 5(2): 51-53, 2007.

[9] Denny Vrandecic, York Sure. "How to Design Better Ontology Metrics". In ESWC 2007, p. 311–325, 2007.