Towards Stable Semantic Ontology Measurement

Yinglong Ma[1,2]

[1] School of Control and Computer Engineering
North China Electric Power University, Beijing 102206, P.R.China
[2] State Key Laboratory of Computer Science
Institute of Software, Chinese Academy of Sciences, Beijing 100190, P.R.China
yinglongma@gmail.com

Abstract. Stable semantic ontology measurement is crucial to obtain significant and comparable measurement results. In this paper, we present a summary of the definition of ontology measurement stability and the preprocessing for stable semantic ontology measurement from [5]. Meanwhile, we describe two existing ontology metrics. For each of them, we compare their stability from the perspectives of structural and semantic ontology measurements, respectively. The experiments show that some structural ontology measurements may be unusable in cases when we want to compare the measurements of different models, unless the pre-processing of the models is performed.

1 Introduction

In recent years, ontology engineers have proposed many ontology metrics for assessing ontology quality such as the literatures [1–4]. However, some proposed ontology metrics are to measure ontology structure instead of ontology semantics which are the nature of ontology. They only simply calculate the number of classes and class inheritances by some labels in ontologies such as `owl:Class` and `rdfs:subClassOf`, and do not consider possibly implicit semantic subsumption between (complex) classes. Most ontology metrics do not take into account the open world assumption (OWA) and the possible addition of implicit axioms, which will cause incomparable measurement results [5]. To use the same metric to measure the ontologies with the same semantic knowledge will bring about variable values. Such ontology metrics may be unstable.

2 Stable ontology measurement and preprocessing

An ontology can be regarded as a set of triples of the form `(s,p,o)`. The structural description of an ontology $\mathcal{O}$ is the set of explicitly represented triples in $\mathcal{O}$. The semantic description of $\mathcal{O}$ is the set that contains not only the structurally described triples, but also all implicit triples obtained by reasoning $\mathcal{O}$. Note that an ontology with the same semantic description possibly has multiple structural descriptions (including $\mathcal{O}$).

Definition 1. *Let $Sem(\mathcal{O})$ be the semantic description of an ontology $\mathcal{O}$. $Sem(\mathcal{O})$ has the multiple structural descriptions, denoted $Stru(\mathcal{O})$=$\{\mathcal{O}, \mathcal{O}_1, \cdots, \mathcal{O}_n\}$. A stable ontology measurement $\mathcal{M}$ is mapping, $\mathcal{M} : Stru(\mathcal{O}) \rightarrow \mathbf{R}$ such that $\mathcal{M}(\mathcal{O}) = \mathcal{M}(\mathcal{O}_1) = \cdots = \mathcal{M}(\mathcal{O}_n)$, where $\mathbf{R}$ is a nonempty set of real numbers.*

We summarize the preprocessing for stable ontology measurement from [5].

1) Naming all anonymous classes and all anonymous individuals. We can automatically detect the related labels and name anonymous classes. Anonymous individuals can be detected and named by class membership. The set of named concepts of Ontology $\mathcal{O}$ is denoted $\mathcal{C}_{\mathcal{O}} = \{C_1, \cdots, C_n\}$, where each C_i is unique, and is either an atomic concept or a named anonymous concept.

2) Eliminating cycles of concept subsumption such as $A \sqsubseteq A_1, \cdots, A_n \sqsubseteq A$, where A, $A_i (1 \leq i \leq n)$ are concepts. Once we detect such a cycle of concept subsumption in an ontology, we replace all cyclic concept subsumption axioms with $B \sqsubseteq A_i$ $(1 \leq i \leq n)$, where B is a new concept name for each cycle.

3) Instances explicitly asserted by class membership, object property and datatype property should be enriched as deeply as possible by reasoning the ontology $\mathcal{O}$.

4) Getting rid of possible transitivity relationships. We attempt to adopt a definition of axiom fanouts per concept to get rid of the possible transitivity relationships. The reason to do this is that a well-founded measurement theory should avoid the *double counting* problem, e.g., a measurement unit is counted more than once. Once some axioms are counted, then the axioms derived from these counted axioms should not be counted. In the following, we specifically discuss axiom fanouts per concept.

Definition 2. *$\forall C, D \in \mathcal{C}_{\mathcal{O}}$, C is directly subsumed by D, i.e., directly-subsumed-by(C, D), iff $\forall C, D \in \mathcal{C}_{\mathcal{O}}(C \sqsubseteq D \wedge \neg\exists C' \in \mathcal{C}_{\mathcal{O}}(C' \sqsubseteq D \wedge C \sqsubseteq C'))$.*

Definition 3. *$\forall C \in \mathcal{C}_{\mathcal{O}}$, the axiom fanouts of C are denoted $AF_C = \{D_1, \cdots, D_m\}$, where for each $D_i (1 \leq i \leq m \leq |\mathcal{C}_{\mathcal{O}}|)$, directly-subsumed-by$(D_i, C)$ holds, and $|\mathcal{C}_{\mathcal{O}}|$ represents the cardinality of $\mathcal{C}_{\mathcal{O}}$.*

In the following, we simply analyze the correction of the preprocessing. On one hand, as mentioned above, for an ontology $\mathcal{O}$, its semantic description $Sem(\mathcal{O})$ contains not only the structural description of $\mathcal{O}$, but also the implicitly expressed knowledge derived from $\mathcal{O}$. This means that, for any axiom or assertion α in $\mathcal{O}$, $\mathcal{O}$ implies α iff $Sem(\mathcal{O})$ implies α. On the other hand, the preprocessing for stable ontology measurement is terminable because Step 1), Step 2) and Step 3) will be terminated if there is no complex concept, cycle of concept subsumption, and unenriched concept in $\mathcal{O}$. At last, this can guarantee that $Sem(\mathcal{O})$ should be finite and unique no matter how the ontology $\mathcal{O}$ is represented. In the case, the measurement result for $\mathcal{O}$ will be invariable and stable if we measure $\mathcal{O}$ by $Sem(\mathcal{O})$. We can also obtain the following corollary.

Corollary 1. *An ontology measurement of ontology's semantic description is stable by using the preprocessing for comparing the measurements of different models.*

3 Proposal of two ontology metrics

A structural ontology measurement is just to measure the explicitly expressed ontology without applying the preprocessing. In contrast to a structural ontology measurement, a semantic ontology measurement is just to measure the quality of semantic description of original ontology by using the preprocessing. We describe two ontology metrics related to axiom fanouts for validating the stability of ontology measurement.

Metric 1: Average Axiom Fanouts per Concept (AAFC)

AAFC of Ontology $\mathcal{O}$ can be defined as follows: AAFC($\mathcal{O}$)=$\frac{\sum_{C\in\mathcal{C}_{\mathcal{O}}} AF_C}{|\mathcal{C}_{\mathcal{O}}|}$.

Metric 2: Average Depth of Concept Subsumption of Leaf Concepts (ADCS-LC)

A concept $RC \in \mathcal{C}_{\mathcal{O}}$ is a root one iff $\neg\exists C \in (\mathcal{C}_{\mathcal{O}} \setminus RC)$ such that $RC \sqsubseteq C$. A concept $LC \in \mathcal{C}_{\mathcal{O}}$ is a leaf one iff $\neg\exists C \in (\mathcal{C}_{\mathcal{O}} \setminus LC)$ such that $C \sqsubseteq LC$. The depth of path p, denoted $|p|$, is the total number of concepts in p. ADCS-LC of $\mathcal{O}$ can be defined as ADCS-LC($\mathcal{O}$)=$\frac{\sum_{p\in PS} |p|}{|LS|}$, where PS and LS are the set of all paths and the set of leaf concepts in ontology $\mathcal{O}$, respectively.

AFC and ADCS-LC are ontology metrics related to ontology fanouts which can be often used as the indicators of some ontology quality properties such as complexity and cohesion [1–4]. They can be used for structural or semantic ontology measurements.

4 Experiments and measurement stability analysis

The goal of the experiments was designed to compare the stability of the two ontology metrics from the perspectives of structural and semantic ontology measurements, respectively. The experimental settings were as follows. 1) We randomly searched the 10 testing ontologies by the search engine, Swoogle. They were evaluated for validating the stability of ontology measurement; and 2) For each of AAFC and ADCS-LC, we collected the values of their structural and semantic ontology measurements, respectively. The measurement values of AAFC and ADCS-LC were shown in Figure 1.

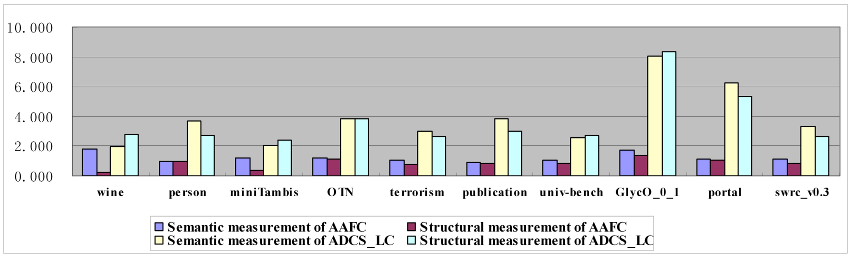

Fig. 1. Semantic and structural measurement values of AAFC and ADCS-LC

Pearson's correlation coefficient is used for analyzing the stability of AAFC and ADCS-LC, which is with the following hypotheses: $H_0 : \rho = 0$ (There is no correlation between the pair of values); $H_1 : \rho \neq 0$ (There is correlation between the pair of values). For each of AAFC and ADCS-LC, we calculated the correlation coefficient and p-value of pair of measurement values. The larger absolute value of the correlation coefficient means stronger correlation between the pair of variables. If the pair of variables are independent, the correlation coefficient is 0. P-values are used in hypothesis tests to either reject or fail to reject a null hypothesis. A small p-value indicates that a null hypothesis is false. A p-value ($<$0.001) means that we must reject the null hypothesis.

By the statistical software, SPSS , we can obtain the correlation coefficient and p-value of pairs of AAFC and ADCS-LC, respectively. We find that the correlation coefficient and p-value between the pair of semantic and structural measurement values of AAFC are -0.195 and 0.588, respectively. This means that there is no obvious correlation between the pair of AAFC. However, there is a very strong correlation between the pair of semantic and structural measurement values of ADCS-LC because their correlation coefficient and p-value are 0.946 and 0.000, respectively. Especially for the p-value, it is less than 0.001 such that we can obviously reject the null hypothesis.

From Corollary 1, if we we want to compare the measurements of different models, a semantic ontology measurement is stable by using the preprocessing. In the case, if the semantic and structural measurements of an ontology metric are strongly correlated, then this means that the structural measurement of the ontology metric may be usable to compare the measurement values of ontologies, and can be a useful indicator of some ontology quality properties such as complexity and cohesion. Otherwise, the structural measurement of the ontology metric is likely to be unusable. According to these analysis, we find that AAFC may be not usable to compare the measurement values of ontologies. In contrast to AAFC, ADCS-LC is usable for both semantic and structural ontology measurement. We believe that more experiments should be made to comprehensively validate the stability of ontology measurement.

5 Conclusions

We summarized the definition about stability of ontology metrics and the preprocessing for stable ontology measurement. By two ontology metrics to compare the measurement stability of different models, we found that some structural ontology measurements may be unusable, unless the pre-processing of the models is performed.

Acknowledgments

This work is supported by National Natural Science Foundation of China (No.61001197), the Fundamental Research Funds for the Central Universities, and State Key Laboratory of Computer Science, IOS, Chinese Academy of Sciences (No.SYSKF1010).

References

1. Gangemi A., Catenacci C., Ciaramita C. and Lehmann J. A theoretical framework for ontology evaluation and validation. *Proceedings of SWAP2005*, 2005.
2. Tartir S., Arpinar I.B., Moore M., Sheth A.P. and Aleman-Meza B. OntoQA: Metric-based ontology quality analysis. *Proceedings of IEEE ICDE 2005 Workshop*, 2005.
3. Orme A., Yao H. and Etzkorn L. Coupling Metrics for Ontology-Based Systems. *IEEE Software*, **23(2)**, (2006) 102–108.
4. Orme A., Yao H. and Etzkorn L. Indicating ontology data quality, stability, and completeness throughout ontology evolution. *Journal of Software Maintenance and Evolution: Research and Experience*, **19(1)**, (2007) 49–75.
5. Vrandecic D. and Sure Y. How to Design Better Ontology Metrics. *Proceedings of ESWC2007*, (2007) 311–325.