Role Of Coupling And Cohesion In Modular Ontologies

Naveen Malik , Sandip Kumar Goyal, Vinisha Malik
Department of Computer Science and Engineering, MM University Mullana , Haryana , India.

Abstract - Ontology is an essential component of the Semantic Web and in recent times, the significance of modular ontologie-s is largely increasing due to their superiority over monolithic ontologies. Out of the multiple modularization choices available, one which guarantees best system design and performance should be applied. Some attributes of modular structure such as cohesion and coupling, determine the goodness of modularization technique applied. Few works are available in the field of assessing modular ontology in terms of cohesion and coupling. We gives an approach for analysing cohesion and coupling of modular ontology. Modular ontology that consists of multiple ontologies (ontology modules). A Good, modular ontology design: "Low coupling, high cohesion" principle. Ontology module coupling: the degree of relatedness between ontology modules low coupling: concepts in a module are not strongly related to concepts in other modules. Ontology module cohesion: the degree of relatedness of classes in a module high cohesion: concepts in a module tend to be strongly related to other concepts in the module.

Keywords- cohesion, coupling , modular ontology

1. INTRODUCTION

ontology is the essential component of the semantic web and nowadays the importance of modular ontology is increasing due to their lead advantage on the monolithic ontologies. Monolithic ontology designs have been used conventionally for the purpose of vocabulary representation. However, monolithic ontology has several shortcomings such as, difficult reuse of concepts and poor scalability, among others. Furthermore, while dealing with large and complex ontologies, the reasoning performance deteriorates. With the expanse of knowledge-base, monolithic ontologies are becoming increasingly difficult to handle. On the other hand, modular ontologies offer numerous benefits, such as partial reuse of components, reduced complexity, customization as per user-requirements, and others. There have been several attempts at modularization of ontologies, yet only few of them are dedicated towards quantification of the quality of modularized structure. This work attempts to analyse the determinants of cohesion and coupling of modularized ontology, which can be instrumental in formulating metrics for the same [2].

Quality assessment of modular ontologies on the basis of cohesion and coupling would assist the end user in making a choice among multiple modular ontologies with varying values of cohesion and coupling. It will not only lead to the selection of the ontology which is logically most consistent, but it will also enhance the system performance.

The rest of the paper is organized as follows. Section 2 gives an Literature review of modular ontology, its development, benefits and challenges. In this section the literature available in this area and explains the importance of cohesion and coupling as tools for measurement of efficiency of modular ontology. This section also highlights the paucity of research work available in this area. Section 3 describes in detail, some open research in the field by citing the limitations of related work. It also emphasizes on the need to devise comprehensive semantic metrics for assessing the quality of modular ontology. Analysis of proposed work is described in section V, which explains our assessment approach for modular ontology, in detail.

2. LITERATURE REVIEW

Authors in [4] propose coupling and cohesion metrics for knowledge-based systems. They assert that coupling exists between two objects if at least one of them affects the other, while cohesion of any object is measured by the level of similarity of its methods. They define frames and measure cohesion (of a frame), and coupling (among multiple frames) using the relationships among the slots of frames.

The work in [2] de¿nes a set of metrics for estimating cohesion and coupling of monolithic and modular ontologies. Authors have employed semantic definitions to categorize various relationships as *strong* and *not-strong (moderate).* However, this metric treats direct and indirect relations in a similar manner and does not incorporate adjustments in metrics owing to the diminishing influence of relationship with increasing distance. Moreover, there are certain relationship types that by definition, do not lie in either *strong* or *moderate* zone. These relationships have not been included in the calculation for cohesion and coupling of the modular ontology.

2.1 Yao et al. [2005] ontology cohesion metrics suite Taxonomy tree considered as the backbone of ontogy.
Number of root classes (NRC), Number of leaf classes (NLC), average depth of leaf nodes in inheritance tree (ADL).Higher values of NRC, NLC and ADL indicate greater separation of concepts.Metrics not suitable for ontology modules

Orme et al. [2006] ontology coupling metrics suite

2.2 NEC (number of external classes used in ontology), REC (number of references to external classes), RI (number of referenced ontologies) (+) Metrics that can be computed locally, without construction of ontology graph

(-) Metrics that cover efferent (aggregation) coupling but not afferent (reuse) coupling. Metrics suite introduced by 2.3 Oh et al. [2011] Two coupling metrics

NSHR, Number of separated hierarchical relations (inter-module links)

NSNR, Number of separated non-hierarchical relations (other inter-module links)

One cohesion metric

Strength of (indirect) connection between two concepts inversely proportional to the length of the shortest path connecting them Cohesion: sum of strength of intra-module links normalized by the number of all possible intra-module links. The main idea of modular ontologies originates from the general notion of modu- lar software in the area of software engineering. Correspondingly, ontology modu- larization can be interpreted as decomposing potentially large and monolithic on- tologies into (a set) of smaller and interlinked components (modules). Therefore, an ontology module can be considered as a loosely coupled and self-contained component of an ontology maintaining relationships to other ontology modules. Thereby, ontology modules are themselves ontologies [4].

2.4 Approaches for Ontology Modularization

In the last years, the problem of ontology modularization attracted more and more attention and, thus, several different approaches for modularizing ontolo- gies appeared. These approaches can be classified in two main categories.

The first main category comprises approaches that focus on the composition of existing ontologies by means of integrating and mapping ontologies. On the one hand, approaches addressing an integration of existing ontologies are OWL Im- port, partial semantic import, package-based description logics, e.g.. On the other hand, mapping approaches basically aim at (inter-)linking sets

3. **ANALYSING OF COUPLING AND COHESION:**

Criteria for Ontology Modularization:

Criteria for modularizing ontologies generally aim at characterising modular on- tologies. To the best of our knowledge only [5] explicitly deals with criteria for ontology modularization. Therefore, [5] distinguish between criteria originating in software engineering, logical criteria, local criteria, structural criteria, quality of modules, and relations between modules. First, criteria from software engineer- ing comprise encapsulation and coherence whereas logical criteria include local correctness and local completeness. Structural criteria, which are also discussed by [6] focus on size and intra-module coherence. It is proposed to determine the quality of modules in terms of module cohesion, richness of the representa- tion, and domain coverage. At least, to assess the relation between modules the criteria connectedness, redundancy, and inter-module distance can be applied. Against this background, the evaluation of ontology modularization respectively applies a subset of the proposed set of criteria with respect to different scenarios and ontology modularization techniques.

Based on best practices in Ontology Engineering, ontology design patterns (ODPs) simplify ontology design by providing “modelling solution to solve recurrent on- tology design problems [1]. Several types of ODPs has already been identified, e.g., logical patterns that are used to solve problems of expressivity, or naming patterns that are conventions for naming elements. Among these types, archi- tectural ontology design patterns (AODPs) aims at describing the overall shape of the ontology.

The strength factor (sf) of any relationship defines the strength of the link between the two concepts involved in that relationship. The influence of both direct and indirect relations can be measured using sf. The sf of a link is calculated using two factors:

1. Type of relation
2 Distance between concepts

The strength of any link can lie in one of the three zones as shown in Figure 1.

a) Strong: These links have a high influence on the cohesion and coupling of the modules. The range of its sf is (2/3, 1].

b) Moderate: These links have a moderate kind of influence on the cohesion and coupling of the modules. The range of its sf is (1/3, 2/3].

c) Weak—These links have a very low influence on the cohesion and coupling of the modules. The range of its sf is (0, 1/3)
A concept holds a dependence on any other concept if the semantics of one concept are influenced by the existence and semantics of the other. Hence, the constraints and attributes of a concept are determined either partially or wholly by the concept it depends on.

Approach that includes :

Domain-independent coupling metrics used in complex network analysis Complex network analysis techniques, hybrid ontology metric set and statistical tests to identify coupling patterns

Graph clustering evaluation metrics as ontology cohesion metrics

Supported by ONGRAM tool :

Integration of OWL2 into SSQSA (Set of Software Quality Static Analyzers) framework

Performs ontology normalization that preserve predefined modularization

Complex class expression represented by one node which references all named concepts contained in the expression

Property domain and property range relations are assembled

Module graph: subgraph of ontology graph induced by ontologies and IMPORTS links

Coupling metrics computed on module graph

Centrality metrics used in complex network analysis

Local: in-, out- and total degree

Global: betweenness, closeness, page rank, etc.

Coupling patterns

Each ontology module characterized by a numeric metric vector M that contains metrics of internal complexity, coupling and cohesion metrics Connected component analysis

Radicchi et al. [2004] notion of community (cluster)

Strong community

$$(\forall i \in C)\ out_{\text{intra-cluster}}(i) > out_{\text{inter-cluster}}(i)$$

Weak community

$$\sum_{i \in C} out_{\text{intra-cluster}}(i) > \sum_{i \in C} out_{\text{inter-cluster}}(i)$$

When a monolithic ontology is modularized, a few links are disconnected due to the modularization. These are the inter-modular links that contribute to the coupling of ontology.

Analogous to cohesion, coupling of module contributed per dependence can be found.

Based on the above inferences, in future, we intend propose precise metrics for cohesion and coupling of modular ontologies. Furthermore, we would do both theoretical and experimental validation of our proposed metrics.

4. APPROACH

In order to characterize reoccurring structures in modularly organized ontologies, the following approach establishes a methodological basis to guide the research programme of this work. This approach comprises six subsequent steps:

1. Search step: the goal of the first step is to gather modularly organized ontologies. Therefore, we use the Semantic Web gateway Watson4 to search for available modular ontologies from the WWW. The search query focused on import-relationships between ontologies covering the same domain. The result is a set of 77 modularly organized ontologies.
2. Cleaning up step: the second step aims at cleaning up the initial search results in order to establish a thorough basis for further experiments. This is necessary because the set of 77 modularly ontologies is afflicted with re- dundancies and incompleteness. As a result, there is a set of 38 modularly organized ontologies constituting a thorough basis for characterizing ontol- ogy modularization.
3. Selection of metrics step: the third step selects a set of appropriate met- rics to characterize modularly organized ontologies. The modular ontologies could be described by various indicators such as the distribution of classes, the network of links between modules, the number of internal links in mod- ules, etc. In general, literature comprises a plethora of various metrics, which could be applied for characterising modular ontologies. As a starting point, this work focuses on metrics originating in the area of software engineering due to its maturity. In particular, this work adopts the following metrics from software engineering to characterize modular ontologies [7]:

* Size of the module: number of classes and properties (object and datatype properties).

* Cohesion of the module: value between 0 and 1 and is specified as follows:

• Hierarchical Class Cohesion (HCC): the number of direct and indi- rect hierarchical class links.

• Role Cohesion (RC): the number of direct and indirect hierarchical role links.

• Object Property Cohesion (OPC): the number of classes which have been associated through the particular object property (domain and range).

* Coupling of the module: it takes an estimation of the inter-dependency of different modules and is specified as follows:

• Hierarchical class dependency (HCD): the number of all direct and indirect hierarchical class relationships to foreign ontologies.

• Hierarchical role dependency (HRD): the number of all direct and indirect hierarchical role relationships to foreign ontologies.

• Object property dependency (OPD): the number of roles that asso- ciate external classes to local ones. • Axiom dependency: a role or a class is associated to an external ontological element through an inclusion axiom

4. Metrics implementation step: the fourth step required to implement met- rics program. The computation was performed by the OWL API5 and the reasoner HermiT6. This step sets up the (technical) evaluation framework.
5. Analysis step: the fifth step is the analysis the basic population of modu- larly organized ontologies.
6. Result step: the sixth step regards the synthesis and discussion of the re- sults from the analysis in order to characterize modular ontologies.

5. Conclusion and Future Work

This work aims at characterizing modularly organized ontologies to contribute to a better understanding of ontology modularization. We introduced the notion of modular ontologies, reported on approaches for ontology modularization, and re- viewed existing efforts to characterize modular ontologies. To characterize mod- ular ontologies, we followed an approach comprising six consecutive steps. This approach mainly includes the extraction and selection of modular ontologies, the selection of a set of metrics from software engineering to analyse modular ontologies, and the evaluation of the analysis results. The evaluation results in a set of four patterns, which allow for characterizing the modularly organization of ontologies. These patterns show amongst others that modularly organized do- main ontologies have a clear structure whereas top-level ontologies tend to have a rather confusing modular organisation.

REFERENCES:

[1] T. Gruber, "Ontology" in the Encyclopedia of Database Systems, L. Liu and M. T. Özsu, Eds. Springer-Verlag, 2009.

[2] Gangemi A. and Presutti V. Ontology Design Patterns, pages 221–243. Springer, Berlin, second edition, 2009.

[3] Frederico Luiz Gonc¸alves de Freitas, Zacharias Candeias Jr., and Heiner Stucken

Towards checking laws' consistency through ontology design: The case of brazilian vehicles' laws. JTAER, 6(1):112–126, 2011.

[4] S. Kramer and H. Kaindl, "Coupling and cohesion metrics for knowledge-based systems using frames and rules," ACM Transactions on Software Eng. and Methodology, pp. 332–358, 2004.

[5] H. Yao, A. M. Orme, and L. Etzkorn, "Cohesion metrics for ontology design and application," J. Comput. Sci., vol. 1, no. 1, pp. 107–113, 2005.

[6] A. M. Orme, H. Tao, and L. H. Etzkorn, "Coupling metrics for ontology-based system," IEEE Software, vol. 23, no. 2, pp. 102–108,2007.

[7] P. S. Doran, V. Tamma, T. R. Payne, and I. Palmisano, "An entropy inspired measure for evaluating ontology modularization," in 5th Int. Conf. on Knowledge Capture, 2009.

[8] Y. Ma, B. Jin, and Y. Feng, "Semantic oriented ontology cohesion metrics for ontology-based systems," J. Syst. and Software, vol. 83, no.1, pp. 143-152, 2010.

[9] Bernardo Cuenca Grau, Bijan Parsia, Evren Sirin, and Aditya Kalyanpur. Mod-ularity and web ontologies. In Patrick Doherty, John Mylopoulos, and Christo- pher A. Welty, editors, Proceedings of KR2006: the 20th International Confer- ence on Principles of Knowledge Representation and Reasoning, Lake District, UK, June 2–5, 2006, pages 198–209. AAAI Press, 2006.

[10] Luciano Serafini and Andrei Tamilin. Drago: Distributed reasoning architecture for the semantic web. In ESWC, pages 361–376. Springer, 2005.

[11] Heiner Stuckenschmidt and Michel Klein. Structure-based partitioning of large concept hierarchies. In International Semantic Web Conference, pages 289–303, 2004.