

Module Based Data Management: A Survey

Manisha P. Bidve¹, N. J. Pathan²

¹SRTM University, Department of CSE, M. S. Bidve Engineering College, Latur, Maharashtra, India,

²D.B.A.T. University, Department of Computer Science and Engineering, Lonere, Raigad, Maharashtra, India,

Abstract: *The new trend for building an ontology based data management system (DMS) is to effort made on design already established a reference model. Use the already existing schemas which are related to the new application need(module) also individualized schemas with extra constraints with respect to new applications and then managing the dataset using these schema. In this survey paper we carry out our investigations in the descriptive logics which underlie modern ontology languages like RDFS, Owl, OWL2 from W3C and DL-lite family which encompasses the base of the QL profile of OWL2, which managing large datasets.*

Keywords: ontologies, Description Logic, OWL

1. Introduction

In many application domain (like medicine or biology) generic schemas resulting from available schemas. Example for this, SOMED is an ontological schema which having more than four lakhs concept, which covers the different medicine, diseases and anatomy (i.e. body) related information as well as the geographical location related to that disease [1].

Mostly ontology engineers takes support by available tools and methodologies which minimize the starting errors. Modularity is useful for the partial reuse of these ontologies. Such as the external ontologies are used as Cystic-Fibrosis and Genetic-Disorder for description of medical research projects [2].

Module extraction problem gives the related endorsement Σ related to terminology \top extract a minimal part \top_1 from \top [3].

Reusability of ontologies: example for this is SNOMED CT (Systematised Nomenclature of Medicine Clinical Terms) is an application of vocabulary often only used a small fraction of its vocabulary of interest. Σ Interpolate instead of whole ontology large ontology are difficult to maintain if making small changes in its axioms it can have effects to that ontology [4].

DL light family, mainly created to capture basic ontology languages while keeping all reasoning task tractable with polynomial time complexity with respect to size of knowledge base. Reasoning means not just computing minor terms between concepts and checking satisfiability of knowledge base but also it is answering complex queries also. DL light family defined by TBOX and ABOX. TBOX gives general properties of concepts and roles. ABOX gives instance (object) of concept and roles. DL light family presents techniques for usual reasoning task of Descriptive Logics. The time complexity of concept and role subsumption, knowledge base satisfiability, and object verification. Time complexity of role and subsumption is polynomial in size of knowledge base [6].

Ontology is common vocabulary for one area of interest. It is easy to build new ontology (reuse of ontology) using available ontology. Newly created large ontology takes some part of the available ontologies involve same concept [7].

Modularity in the context of *collaborative ontology development* and *controlled integration* and defined what it means for an ontology we are developing to be safely integrated with a external ontology[9].

To delete some terms from the ontology is difficult task because it generally affects the relationship between remaining terms [10].

2. Related Work

The ontology defines different types of projects which focuses on a specific topic. Consider an ontology engineer defines concept like Genetic_Disorder_Project and Cystic_Fibrosis_EUProjects in his ontology p. Genetic_Disorder_Project describes about genetic disorder and other one describes about European Projects about cyclic fibrosis given by p1 and p2 in table. The ontology engineer is an expert on research projects and then he decides to use knowledge about these subjects repeatedly from mostly used and well established medical ontologies. Most simple method to reuse these concepts to import foreign medical ontologies. It may uses (import) a module instead of fully ontology. The goal is to accessibility the visualization of navigation through the ontology. The algorithm uses set of heuristic for calculating the fragments of dependency between concepts in ontology and output a graphical representation of these dependencies.[2]

Table 1: Comparison of Different Modularization Algorithms

Ontology	NCI	SNOMED	GO	SUMO
Atomic concepts	27772	255318	22357	869
AI: Prompt-Factor	Max.(%) 87.6	100	1	100
	Max.(%) 75.84	100	0.1	100
AI: Prompt-Factor	Max.(%) 55	100	1	100
	Max.(%) 30.8	100	0.1	100
AI: Prompt-Factor[2]	Max.(%) 0.8	0.5	0.4	2
	Max.(%) 0.08	0.05	0.05	0.09

For evaluation and comparison author implemented some algorithms.

- A1:-The PROMPT_FACTOR algorithm
- A2:-The algorithm for extracting modules
- A3:-The algorithm for extracting module based on syntactic locality

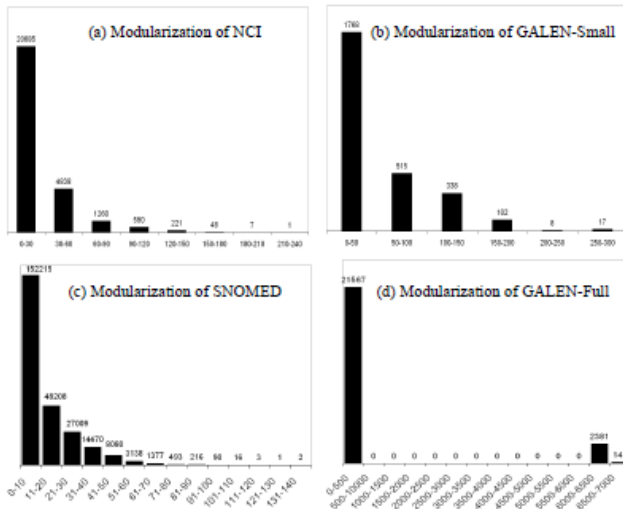


Figure 1: Distribution for the sizes of syntactic locality-based modules for atomic concepts: the X-axis gives the number of concepts in the modules and the Y-axis the number of modules for each size range.

From [3], when two medical terminologies are semantically inseparable, then it does not make any difference.

$$\Sigma = \{A, B\} \text{ and } T1 = \{A \sqsubseteq \exists r.B\}$$

It is computationally difficult to finding semantic notions of entailment and inseparability instead of using a reduction of validity of bimodal formula frame.

2.1 DL-Lite Family

From [4], if is all its concepts inclusion and equations are of the form $A \sqsubseteq C$ and $A \equiv C$ and no concept name more than once on the left hand side then use $A \boxtimes C$ to denote expression of form $A \sqsubseteq C$ and $A \equiv C$. Collectively patients contents into computational resources requires the use of formal as well as logical representation of anatomical knowledge. The field of computing science concentrated on structuring knowledge is called "ontology". This ontology knowledge of anatomical knowledge performed with very high level detailed information and complexity FMA ontology is used for scientific community in today for specific community. FMA is open source ontology used for general use. It is a structure related to human body[5]. Descriptive Logics (DLs) that represent the domain of interest in terms of concept, denoting sets of objects and roles denotes binary relation between concepts. Different DLs allows for the different constructs [6]. The description logic (DLs) is OWL DL which is modern ontology language. The syntax of this DL is given by the signature S and set of constructors. A signature S is the union of set C of atomic concept(A,B,...) which represents set of elements and a set R of atomic roles(r,s,...) represents binary relation between elements[9].

3. Conclusion

The modules explained in paper generalize both modules obtained by extracting a subset of Tbox with respect to selected relation [4] or by forgetting about relations. And with existing work, there may be problem of safe individualized module built from existing reference DMS. This occurs new issues to check simply that a module based DMS generate independently but compatible with respect to the reference DMS from which it has been built. [6] Extract modules from DL-lite schema following a forgetting approach.

References

- [1] Modular Ontologies: Concepts, Theories And Techniques For Knowledge Modularization, H. Stuckenschmidt, C. Parent, S. Spaccapietra, eds. Springer, 2009.
- [2] B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler, "Extracting Modules From Ontologies: A Logic-Based Approach," Proc. Third Int'l Workshop OWL Experiences and Directions (OWLED), 2007.
- [3] B. Konev, C. Lutz, D. Walther, and F. Wolter, "Semantic Modularity and Module Extraction in Description Logics," Proc. 18th European Conf. Artificial Intelligence (ECAI), 2008.
- [4] B. Konev, D. Walther, and F. Wolter, "Forgetting and Uniform Interpolation in Large-Scale Description Logic Terminologies," Proc. 21st Int'l Joint Conf. Artificial intelligence (IJCAI), 2009.
- [5] O. Palombi, G. Bousquet, D. Jospin, S. Hassan, L. Revéret, and F. Faure, "My Corporis Fabrica: A Unified Ontological, Geometrical And Mechanical View Of Human Anatomy," Proc. Second Workshop 3D Physiological Human (3DPH), 2009.
- [6] D. Calvanese, G.D. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati, "Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family," J. Automated Reasoning, vol. 39, no. 3, pp. 385-429, 2007.
- [7] R. Kontchakov, L. Pulina, U. Sattler, T. Schneider, P. Selmer, F. Wolter, and M. Zakharyashev, "Minimal Module Extraction from DL-Lite Ontologies Using QBF Solvers," Proc. 21st Int'l Joint Conf. Artificial Intelligence (IJCAI), 2009.
- [8] S. Ghilardi, C. Lutz, and F. Wolter, "Did I Damage My Ontology? A Case for Conservative Extensions in Description Logics," Proc. 10th Int'l Conf. Principles of Knowledge Representation and Reasoning (KR), 2006.
- [9] B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler, "Just the Right Amount: Extracting Modules from Ontologies," Proc. 16th Int'l Conf. World Wide Web (WWW), 2007.
- [10] Z. Wang, K. Wang, R.W. Topor, and J.Z. Pan, "Forgetting concepts in DL-Lite," Proc. Fifth European Semantic Web Conf. Semantic Web: Research and Applications (ESWC), 2008.

Author Profile



M. P. Bidve received the B.E. degree in Computer Science and Engineering from M.S. Bidve Engineering College in 2013. Now, she is pursuing Master's in Engineering (Computer Science and Engineering) from M.S. Bidve Engineering college, Latur , SRTM University Nanded, Maharashtra.



N. J. Pathan received the B.E degree in CE and T from S. G. G. S. College of Engineering, Nanded in 1991, and M.Tech. degree in Computer Science & Engineering from D. B. A. T. U. A College of Engineering, Lonere , Raigad in 2006. He is in teaching profession since 1991 and now he is with M.S. Bidve Engineering College, Latur (M.S.) as Associate Professor since 2006.