

A Contextualized Knowledge Framework for Semantic Web

Mathew Joseph^{1,2}

¹ FBK-IRST, Via Sommarive 18, 38050 Trento, Italy

² DISI, University of Trento, Italy
mathew@fbk.eu

Abstract. This thesis focuses on developing an efficient framework for contextualized knowledge representation on Semantic Web. We point out the drawbacks of existing formalism for contexts that hinder an efficient implementation and propose a context formalism that enables the development of a framework with desired properties. Some of the future milestones for this thesis work are to (i) develop a proof theory for the logical framework based on Description Logics (DL) (ii) develop reasoning algorithms (iii) verify and compare the performance of these algorithms to existing distributed reasoning formalisms and (iv) implement the system.

1 Problem, Motivation and Related Work

Contexts[1] have been appreciated for their ability to use the “Divide and Conquer” strategy for problem solving. They help to localize reasoning[2] and search tasks, allow us to represent conflicting information by separating them in different contexts. Moreover they are a potential hope for the current state-of-the-art semantic web reasoners and ontology development tools for handling today’s large Knowledge Bases (KB) by allowing to partition these KBs to distributed smaller ones.

In spite of past initiatives for contextualized knowledge representation[3,4] not enough focus has been applied on producing a computationally tractable framework. As far we know not a single approach has been implemented to demonstrate its fruitfulness on the semantic web.

Also the initiatives in the past to provide a logical framework of contexts, many have been propositional in nature[5,6,7] and the other expressive versions based on *quantificational logic* such as the one given in [3] do not provide a sound and complete proof theory. [8] is based on the assumptions such as *rigid identifier property*¹, *barcan formula compliance*² and does not have a decidable logic. Also complications due to arbitrary nesting of contexts and lack of efficient decision procedures are some of the reasons for unsuitability of these frameworks on the Semantic web.

Some of the inherent problems in the works done so far is that contexts have been treated as first-class objects [3,8], allowing to use any fragment of first order logic on these objects. But how a relation $R(C_1, C_2)$ between two context objects C_1 and C_2 reflects reasoning in the knowledge part of these contexts is unspecified.

¹ Same URI denotes the same entity every where.

² Interpretation domains are the same for every context.

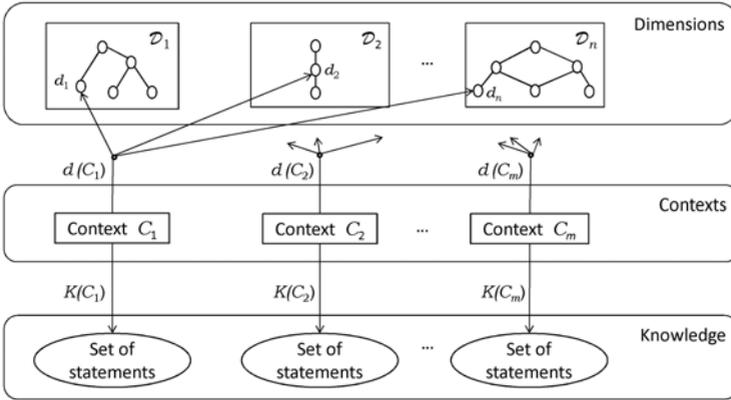


Fig. 1. Contextualized knowledge repository

2 Proposed Approach

A schematic diagram for a Contextualized knowledge repository is given in Fig. 1. Inspired by [2,9], we also comply to the fact that contexts have a logical theory and a fixed set of dimension values. These dimension values (time, location, topic etc.) can be used to qualify the contexts and are thus meta-knowledge for the contexts, hence the logical formalism employed for representation should allow for this meta-knowledge description and use this description for reasoning.

For reasons of tractability and applicability of our logical framework on the semantic web we ground our approach on OWL2 DL. We suppose that this can be extended to full first order logic. We arrive at the following definitions:

Definition 1 (Contextual dimension). A context dimension, \mathcal{D} is an OWL2 knowledge base whose signature $\Sigma(\mathcal{D})$ contains a set of constant symbols, D , called dimension values, and a strict partially ordered binary relation \prec called coverage relation, which is the only relation present³

Definition 2 (Context). Given a set of n dimensions $\{\mathcal{D}_i\}_{i=1..n}$, and a signature Σ , a context \mathcal{C} in the dimensions $\{\mathcal{D}_i\}_{i=1..n}$ is a pair $\langle \mathbf{d}(\mathcal{C}), K(\mathcal{C}) \rangle$, where

1. $\mathbf{d}(\mathcal{C}) = \langle d_1, \dots, d_n \rangle \in D_1 \times \dots \times D_n$;
2. $K(\mathcal{C})$ is an OWL2 knowledge base on the signature Σ ;

We define the knowledge about the context objects in *context description space* and restrict its logic to contain only one predicate \prec with a well defined semantics to relate any two context objects⁴. This is because we currently restrict our focus on qualifying context objects with dimensions – location, time and topic, where we use the \prec to

³ Strict partially ordered relations can be represented in OWL2.

⁴ This can be extended to a fixed number of predicates with well defined semantics.

represent geographic containment, temporal containment and topic coverage relations respectively⁵.

Definition 3 (Context Description Space). A context description space, \mathfrak{K} is an OWL2 knowledge base whose signature $\Sigma(\mathfrak{K})$ contains a set of context objects \mathcal{C}_i , a set of functional symbols $\mathcal{D}_i\text{-Of}(\cdot)$, the binary relation \prec called context cover relation, which is the only binary relation on context objects.

Intuitively the $\mathcal{D}_i\text{-Of}(\cdot)$ is a function to associate a context \mathcal{C} to value of \mathcal{D}_i , context-cover relation \prec is such that for any two contexts \mathcal{C}_a and \mathcal{C}_b , $\mathcal{C}_a \prec \mathcal{C}_b$ iff $\mathcal{D}_i\text{-Of}(\mathcal{C}_a) \prec_i \mathcal{D}_i\text{-Of}(\mathcal{C}_b)$ for $i = 1, \dots, n$.

Definition 4 (Contextualized Knowledge Base). Given a signature Σ , a contextualized knowledge base (CKB) is a tuple $\mathfrak{K} = (\mathcal{D}, \mathfrak{K}, \mathcal{C})$ where

1. $\mathcal{D} = \langle \mathcal{D}_1, \dots, \mathcal{D}_n \rangle$ is an n -tuple of context dimensions
2. \mathfrak{K} is the outer context
3. \mathcal{C} is a set of contexts on dimensions $\mathcal{D}_1, \dots, \mathcal{D}_n$

The relation \prec between any two context objects \mathcal{C}_1 and \mathcal{C}_2 describe the relation between domains of these contexts. For example if $\mathcal{C}_1 \prec \mathcal{C}_2$ then, (i) domain of \mathcal{C}_1 is contained in the domain of \mathcal{C}_2 (ii) identifiers in \mathcal{C}_1 and \mathcal{C}_2 are rigid. We strongly identify the need of such a relation \prec with such special semantics for contexts, because in the semantic web we are normally dealing with URIs and these normally retain their meaning across contexts.

Also note the use of meta-knowledge in our framework, such meta-knowledge values such as time, location, topic or provenance can reveal significant facts about contexts. For example if it is the case that all the dimension values for two Contexts \mathcal{C}_1 and \mathcal{C}_2 are respectively equal, then $\mathcal{C}_1 \prec \mathcal{C}_2$ and $\mathcal{C}_2 \prec \mathcal{C}_1$, this means that information in these contexts are highly coherent and there can be much greater extent of lifting of axioms and migration of individuals between them.

Example 1. Suppose \mathcal{C}_{greece} and \mathcal{C}_{europe} are two contexts in our context description space and $\mathcal{C}_{greece} \prec \mathcal{C}_{europe}$, then for a formula such as

$$\mathcal{C}_{greece} : \text{Conference}(\text{eswc-2010})$$

it is also the case that

$$\mathcal{C}_{europe} : \text{Conference}_{greece}(\text{eswc-2010})$$

In this case the assumption is that domain of discourse of \mathcal{C}_{greece} is covered by that of \mathcal{C}_{europe} . The constant symbol **eswc-2010** means the same across these contexts and extension of concepts **Conference** and **Conference_{greece}** are equal. such lifting of axioms across contexts can be taken care by rules of the system such as

$$\frac{\mathcal{C} : \Phi \quad \mathfrak{K} : \mathcal{C} \prec \mathcal{C}'}{\mathcal{C}' : \Phi'}$$

where \mathfrak{K} is the context description space and Φ' and Φ are formulas

⁵ This is a philosophically debatable assumption, but in general setting a fixed set of dimensions can completely qualify the context space[9].

3 Methodology

Problem definition. The problem is to develop a usable framework for context dependent knowledge representation for the semantic web. The framework developed should support the growing semantic web needs like distributed reasoning, tractability, scalability, reasoning on meta-knowledge for deriving facts in the knowledge.

Survey of existing approaches. A description of alternate approaches on contexts that has been considered is in Section 1. We also consider existing works on distributed reasoning like DDL[10,11] and collaborate with their inventors for reusing existing results and implementations. Some of the shortcomings of these existing frameworks and the ones in [12,13], are that they do not take into consideration, the use of meta-knowledge. Although approaches like [14] does reasoning on meta-knowledge, their semantics do not suffice to relate two distinct pieces of knowledge and how axioms can be lifted across them.

Design of Initial framework and Prototype Implementation. A brief description of the designed framework has been given in Section 2, We are currently working on specifications of syntax and semantics of our framework. The current problem we are addressing is the semantics of *lifting rules*⁶. Currently we limit our selves to simple syntax of lifting rules. A Prototype Implementation has been constructed for testing and for running future experiments.

Development of proof theory and reasoning algorithms. The next immediate step to be considered is development of a proof theory that is sound and complete with respect to semantics employed. For every DL axiom Φ , We have an associated context C , and hence we denote this fact by $C : \Phi$. Then we need to consider development of sound and complete reasoning algorithms.

System Implementation and Testing. Proposed framework once designed need to be implemented to test it's ability of applicability on semantic web, verify correctness of implemented reasoning algorithms.

4 Results and Status

We have implemented a first prototype of a contextualized knowledge repository, by restricting the expressivity of data in the repository to RDF(S).⁷ Such a restriction is explained by the scalability issues of available tools. Practically, we grounded our prototype on Sesame RDF(S) storage and querying framework⁸. We also demonstrated with this prototype application how search efficiency and accuracy on semantic web can be benefited with our notion of contexts[15]. In this work we demonstrated the use of context dimensions to form a hierarchical structure of contexts, which can guide the order of search. Another application of contexts for semantic enrichment was demonstrated in [16]. We are currently in the phase of development of the required proof theory.⁹

⁶ The rules that relate information in different contexts.

⁷ The prototype is available at <https://dkm.fbk.eu/index.php/CKR>

⁸ <http://www.openrdf.org/>

⁹ I am a first year PhD Student.

5 Conclusions and Future Work

This paper introduced the current problems of semantic web, that motivates the idea of a contextualized knowledge representation framework. We explained the related work done before on contexts. We introduced our approach (section 2) as a solution and stated the methodology of research adopted during this thesis (section 3) and results obtained.

Some of the future milestones to be achieved for this thesis work are (i) develop a proof theory for the logical framework based on Description Logics (DL) (ii) developing reasoning algorithms (iii) verify and compare the performance of these algorithms to existing distributed reasoning formalisms and (iv) system implementation.

References

1. McCarthy, J.: Generality in Artificial Intelligence. *Communications of the ACM* 30(12), 1029–1035 (1987)
2. Benerecetti, M., Bouquet, P., Ghidini, C.: Contextual Reasoning Distilled. *Experimental and Theoretical AI* 12(3), 279–305 (2000)
3. Guha, R.: Contexts: a Formalization and some Applications. PhD thesis, Stanford (1992)
4. Guha, R., McCool, R., Fikes, R.: Contexts for the semantic web. In: McIlraith, S.A., Plexousakis, D., van Harmelen, F. (eds.) *ISWC 2004*. LNCS, vol. 3298, pp. 32–46. Springer, Heidelberg (2004)
5. Buvac, S., Mason, I.A.: Propositional logic of context. In: *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pp. 412–419 (1993)
6. Giunchiglia, F., Serafini, L.: Multilanguage hierarchical logics, or: how we can do without modal logics. *Artificial Intelligence* 65(1), 29–70 (1994)
7. Ghidini, C., Giunchiglia, F.: Local models semantics, or contextual reasoning = locality + compatibility. *Artificial Intelligence* 127 (2001)
8. Buvac, S.: Quantificational logic of context. In: *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pp. 412–419 (1996)
9. Lenat, D.: The Dimensions of Context Space. Tech. Rep., CYCorp (1998), <http://www.cyc.com/doc/context-space.pdf>
10. Borgida, A., Serafini, L.: Distributed description logics: Assimilating information from peer sources. *J. Data Semantics* 1, 153–184 (2003)
11. Serafini, L., Taminin, A.: Drago: Distributed reasoning architecture for the semantic web. In: Gómez-Pérez, A., Euzenat, J. (eds.) *ESWC 2005*. LNCS, vol. 3532, pp. 361–376. Springer, Heidelberg (2005)
12. Bouquet, P., Giunchiglia, F., van Harmelen, F., Serafini, L., Stuckenschmidt, H.: C-owl: Contextualizing ontologies. In: Fensel, D., Sycara, K., Mylopoulos, J. (eds.) *ISWC 2003*. LNCS, vol. 2870, pp. 164–179. Springer, Heidelberg (2003)
13. Kutz, O., Lutz, C., Wolter, F., Zakharyashev, M.: E-connections of abstract description systems. *Artificial Intelligence* 156(1), 1–73 (2004)
14. Schenk, S., Dividino, R.Q., Staab, S.: Reasoning with provenance, trust and all that other meta knowlege in owl. In: *SWPM* (2008)
15. Joseph, M., Serafini, L., Taminin, A.: Context shifting for effective search over large knowledge bases. In: *Workshop on Context, Information And Ontologies (CIAO 2009)*, collocated with the 6th European Semantic Web Conference, *ESWC 2009* (2009)
16. Taminin, A., Magnini, B., Serafini, L., Girardi, C., Joseph, M., Zanolini, R.: Context-driven semantic enrichment of italian news archive. In: *ESWC 2010* (2010) (accepted for publication)